OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 03:29:26

```
Title:          US-09-996-617-2
Perfect score:  7534
Sequence:       1 MAGGAWGRLACYLBFLKKEE.........HLIMELWEKGSKKGLLPLSS 1429

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:   1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  A_Geneseq 19Jun03:*
   1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
   2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
   3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
   4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
   5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
   6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
   7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
   8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
   9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
  10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
  11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
  12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
  13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
  14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
  15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
  16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
  17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
  18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
  19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
  20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
  21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
  22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
  23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
  24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

; Search time 69.5142 seconds
(without alignments)
3262.936 Million cell updates/sec
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7534	100.0	1429	22	AAB62571	Human CARD-7 polyp
2	7534	100.0	1429	23	ABG97969	Human leucine rich
3	7534	100.0	1429	23	ABG78472	Leucine-rich repea
4	7534	100.0	1429	23	ABB77916	Human leucine-rich
5	7534	100.0	1429	24	ABG71631	Human caspase recr
6	7534	100.0	1429	24	ABG71633	Human caspase recr
7	7502	99.6	1473	22	AAY72669	Human NB-ARC and C
8	7502	99.6	1473	22	AAY72671	Human NAC beta iso
9	7488	99.4	1429	23	ABG78455	Human NAC caspase recr
10	7488	99.4	1429	23	AAO17855	Pyrin domain conta
11	7484	99.3	1473	22	AAB06758	Human G-protein co
12	7342	97.5	1397	22	AAY72670	Human NB-ARC and C
13	7325.5	97.2	1442	22	AAY72671	Human NB-ARC and C
14	6344	84.2	1454	22	AAY72673	Human NAC beta iso
15	6174	81.9	1424	21	AAY72674	Human NAC gamma or
16	2171.5	28.8	442	21	AAB24513	Human secreted pro
17	1380	18.3	1061	23	AAO15590	Pyrin domain conta
18	1375	18.3	1099	23	AAO17857	Pyrin domain conta
19	1279	17.0	1034	22	AAE07514	Huma PYRIN-8 prot
20	1279	17.0	1034	24	ABU08503	Human PYRIN-1 prote
21	1263.5	16.8	920	23	ABP53254	Human MDDT-13 prot
22	1154	15.3	896	23	AAO15592	Human PYRIN-11 pro
23	1037	13.8	994	23	AAO15593	Human PYRIN-3 prot
24	1020.5	13.6	983	23	AAO17870	Pyrin domain conta
25	1007	13.4	919	24	ABP43483	Human secreted pro
26	1007	13.4	919	24	ABJ19319	Human PYRIN-1 prote
27	985.5	13.1	1162	23	AAE25053	Human EMBRY-1 prot
28	983	13.0	1200	23	AAO17856	Pyrin domain conta
29	983	13.0	1200	24	AAU79526	Human MATER protei
30	982.5	13.0	2312	24	AAE31749	Human MATER protei
31	972.5	12.9	858	23	AAO17859	Pyrin domain conta
32	948	12.6	1062	22	AAM50328	NOVX related prote
33	948	12.6	1062	23	AAG65895	Human secreted pro
34	945.5	12.5	1062	23	AAO17856	Pyrin domain conta
35	945.5	12.5	1033	22	AAE07513	Pyrin domain conta
36	945.5	12.5	1033	22	AAG78456	Amino acid sequenc
37	945.5	12.5	1344	24	AAU08456	Human nucleotide b
38	941.5	12.5	1049	22	AAU08502	Human nucleotide b
39	934.5	12.4	1033	23	AAO17856	Amino acid sequenc
40	924.5	12.3	732	23	AAO15585	Human pyrin domain
41	922.5	12.2	952	22	AAG65894	Human nucleotide b
42	910	12.1	1022	23	AAU65214	Human nucleotide b
43	908.5	12.1	1022	23	AAO17858	Human PYRIN-5 prot
44	908.5	12.1	980	23	AAG03924	Amino acid sequenc
45	897	11.9	891	22	AAO17862	Human PYRIN-5 prot

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
AAB62571

ID   AAB62571 standard; Protein; 1429 AA.
XX
AC   AAB62571;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Human CARD-7 polypeptide.
XX
KW   CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;
KW   autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;
KW   antibacterial; antiviral; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FH
FT   Domain           329..645
FT                    /note= "nucleotide binding domain"
FT   Domain           333..341
FT                    /note= "kinase 1A subdomain"
FT   Domain           404..413
FT                    /note= "kinase 2 subdomain"
FT   Domain           454..463
FT                    /note= "kinase 3a subdomain"
FT   Domain           615..622
FT                    /note= "motif 2 domain"
FT   Domain           807..834
FT                    /note= "leucine-rich domain"
FT   Domain           836..863
FT                    /note= "leucine-rich domain"
FT   Domain           864..891
FT                    /note= "leucine-rich domain"
```

Pyrin domain conta
Human G-protein co
Human NB-ARC and C
Human NB-ARC and C
Human NAC beta iso
Human NAC gamma or
Human secreted pro
Pyrin domain conta
Pyrin domain conta
Huma PYRIN-8 prot
Human PYRIN-1 prote
Human MDDT-13 prot
Human PYRIN-11 pro
Human PYRIN-3 prot
Pyrin domain conta
Human secreted pro
Human PYRIN-1 prote
Human EMBRY-1 prot
Human MATER protei
Human MATER protei
Pyrin domain conta
NOVX related prote
Human secreted pro
Pyrin domain conta
Pyrin domain conta
Amino acid sequenc
Human nucleotide b
Human nucleotide b
Amino acid sequenc
Human pyrin domain
Human nucleotide b
Human nucleotide b
Human PYRIN-5 prot
Amino acid sequenc
Human PYRIN-5 prot
Human NOV125b prot
Novel human diagno
Pyrin domain conta
Pyrin domain conta
Human CGDD SEQ ID
Human nucleotide b
Human nucleotide b
Human G-protein co

```
FT   Domain            893..920
FT                     /note= "leucine-rich domain"
FT   Domain            921..948
FT                     /note= "leucine-rich domain"
FT   Domain            950..976
FT                     /note= "leucine-rich domain"
FT   Domain
FT                     /note= "leucine-rich domain"
XX
PN   WO200130813-A1.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-US29796.
PR   27-OCT-1999; 99US-0428252.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI; 2001-343340/36.
DR   N-PSDB; AAF83651.
XX
PS   Disclosure; Fig 1A-D; 80pp; English.
XX
PT   Identifying a modulator of interaction between caspase recruitment
PT   domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
PT   comprises measuring the binding of CARD-7 and CARD-5 in the presence of
PT   the compound -
XX
CC   The invention relates to identifying compounds, that modulate interaction
CC   between caspase recruitment domain (CARD)-7 and CARD-5. The method
CC   involves measuring the binding of CARD-7 and CARD-5 in the presence of
CC   the compound (an increase in the binding of CARD-7 to CARD-5 in the
CC   presence of the compound compared to the binding in the absence of the
CC   compound indicates that the compound is a modulator of CARD-7-CARD-5
CC   interaction). Modulators of CARD-7 and CARD-8 expression or activity can
CC   be used to treat or diagnose disorders such as cancer, bacterial or
CC   viral infections, autoimmune disorders (systemic lupus erythematosus,
CC   immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
CC   organ-specific autoimmunity including multiple sclerosis, Hashimoto's
CC   thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
CC   CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
CC   of cellular processes including cell growth and cell death. The present
CC   sequence represents the human CARD-7, an intracellular protein.
XX
SQ   Sequence 1429 AA;

Query Match             100.0%;  Score 7534;  DB 22;  Length 1429;
Best Local Similarity   100.0%;
Matches  1429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy         61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSPFYSPSEPHLGSPSQPTSTAVLMPW 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSPFYSPSEPHLGSPSQPTSTAVLMPW 120

Qy        121 IHELPAGCTQGSERRVLRQLPDTSGRRWRBISASLLYQALPSSSDPHESPSQESPNAPTST 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IHELPAGCTQGSERRVLRQLPDTSGRRWRBISASLLYQALPSSSDPHESPSQESPNAPTST 180

Qy        181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERBREREKSEKGRPPWAAVGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERBREREKSEKGRPPWAAVGT 240

Qy        241 PPQAHTSLQPHHHWPEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPQAHTSLQPHHHWPEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy        301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGBAGIGKSTLARQVKEAWGRGQLYGDR 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGBAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFPILDGVDEPGWVLQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFPILDGVDEPGWVLQ 420

Qy        421 EPSSELCLHWSQPDPADALLGSLLGKTILPEASFLITARTTALQNIIPSLSEQARWEVLG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EPSSELCLHWSQPDPADALLGSLLGKTILPEASFLITARTTALQNIIPSLSEQARWEVLG 480

Qy        481 FSESSRKEYFYVRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKBKL 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FSESSRKEYFYVRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKBKL 540

Qy        541 TLISKTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TLISKTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600

Qy        601 ISTFLKMGILQEHPIPLSYSPIHLCPQEFFAAMSVYLBDEKGRGKHSNCIIDLLEKTLEAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ISTFLKMGILQEHPIPLSYSPIHLCPQEFFAAMSVYLBDEKGRGKHSNCIIDLLEKTLEAY 660

Qy        661 GIHGLFGASTTRFLLGLLSDBDGEREMENIFHCRLSQGGRNLMQWVPSLQLILLQPHSLESLH 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GIHGLFGASTTRFLLGLLSDBDGEREMENIFHCRLSQGGRNLMQWVPSLQLILLQPHSLESLH 720

Qy        721 CLYETRNKTFLITQVWAHFEEMGMCVETDMELLVCTFPCIKFSRHVKKLQLIEGRQHRSTWS 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CLYETRNKTFLITQVWAHFEEMGMCVETDMELLVCTFPCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy        781 PTMVVLFRWVFVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PTMVVLFRWVFVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840

Qy        841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHICQRLRQPSCKLQRLQ 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHICQRLRQPSCKLQRLQ 900

Qy        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy        961 TTLSDEMRQELRALEQEBKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTLSDEMRQELRALEQEBKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy       1021 ASHVAQANLKLLDVSKIFPIAEIAEESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ASHVAQANLKLLDVSKIFPIAEIAEESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy       1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140

Qy       1141 PQHSWMVAGPLLDIKABPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 PQHSWMVAGPLLDIKABPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200

Qy       1201 ELHHIVLENPSFSPLGVLLKMIHHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ELHHIVLENPSFSPLGVLLKMIHHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1260

Qy       1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320

Qy       1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy       1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKGSKKGLIPLSS 1429
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKGSKKGLIPLSS 1429
```

```
Db
      RESULT 2
      ABG97969
ID    ABG97969 standard; Protein; 1429 AA.
XX
AC    ABG97969;
XX
DT    07-JAN-2003 (first entry)
XX
DE    Human leucine rich repeat domain containing protein #2.
XX
KW    Leucine rich repeat; nervous system; human; neural disorder; apoptosis;
KW    renal disorder; immune disorder; arthritis; asthma; AIDS;
KW    acquired immunodeficiency syndrome; rheumatoid arthritis;
KW    haematopoietic disorder; metabolic disorder; reproductive disorder;
KW    pulmonary disease; cardiovascular disease; hyperproliferative disorder;
KW    neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW    Huntington's disease; developmental disorder; autoimmune disease;
KW    Addison's disease; haemolytic anaemia; antiphospholipid syndrome;
KW    allergic encephalomyelitis; gene therapy.
XX
OS    Homo sapiens.
XX
XX
PN    WO200274959-A2.
XX
PD    26-SEP-2002.
XX
PF    20-DEC-2001; 2001WO-US50457.
PR    03-JAN-2001; 2001US-259479P.
PR    09-JAN-2001; 2001US-260616P.
XX
PA    (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI    Ramanathan C, Feder J, Mintier G;
XX
PT    New HLRRNS1 nucleic acids and polypeptides, useful for preventing,
PT    treating, or ameliorating e.g. renal disorder, immune, hematopoietic,
PT    metabolic, reproductive, pulmonary, cardiovascular or autoimmune
PT    diseases -
XX
PS    Example 1; Page 392-397; 415pp; English.
XX
CC    The invention describes nucleic acids encoding human leucine-rich repeat
CC    containing proteins expressed in nervous system tissues, HLRRNS1. The
CC    HLRRNS1 polypeptide or the polynucleotide is useful for preventing,
CC    treating, or ameliorating a neural disorder or a disorder related to
CC    aberrant apoptosis modulation (either directly or indirectly), renal
CC    disorder, immune disorder (e.g. arthritis, asthma, acquired
CC    immunodeficiency syndrome (AIDS) or rheumatoid arthritis),
CC    haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
CC    diseases, hyperproliferative disorders, neurodegenerative diseases
CC    (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
CC    disease), developmental disorders, non-infectious disorders, nervous
CC    system diseases and/or disorders, and autoimmune diseases (e.g.
CC    Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or
CC    allergic encephalomyelitis). The polynucleotides are also useful as
CC    chromosome markers, for chromosome identification, gene therapy, and in
CC    identifying organisms from minute biological samples. This is the amino
CC    acid sequence of a leucine-rich repeat containing protein.
XX
SQ    Sequence   1429 AA;

Qy    1 MAGGAWGRLACYLEFLKKEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Query Match        100.0%;   Score 7534;   DB 23;   Length 1429;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1429;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;
```

```
Qy    1 MAGGAWGRLACYLEFLKKEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAGGAWGRLACYLEFLKKEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy  121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQOESPNAPTST 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQOESPNAPTST 180

Qy  181 AVLGSWGSPSPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERERESEKGRPPWAAVVGT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AVLGSWGSPSPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERERESEKGRPPWAAVVGT 240

Qy  241 PPQAHTSLQPHHIHPWEPSVRESLCSTWPKWXNEDFPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 PPQAHTSLQPHHIHPWEPSVRESLCSTWPKWXNEDFPNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy  301 PDYVEENRGHLIEIRDLFGPGLDTQEBPRIVILQGRAAGIGKSTLARQVKEAWGRGQLYGDR 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy  361 FQHVPYFSCRELAQSKVVSLAELLGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FQHVPYFSCRELAQSKVVSLAELLGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy  421 EPSSELCLHWSQOPPADALLGSLLGKTILPEASFLITARTTALONLIPSLEQARWVEVLG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 EPSSELCLHWSQOPPADALLGSLLGKTILPEASFLITARTTALONLIPSLEQARWVEVLG 480

Qy  481 FSESSRKEYFYRYFFDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 FSESSRKEYFYRYFFDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy  541 TLTSKTTTTLCIHYLAQALQAQPLGPOLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TLTSKTTTTLCIHYLAQALQAQPLGPOLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy  601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy  661 GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy  781 PTMVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 PTMVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy  961 TTLSDEMRQELRALEQEKFQLLIFSRRKFSVMTPTEGLDTGEMSNSTSSSLKRQRLGSERA 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TTLSDEMRQELRALEQEKFQLLIFSRRKFSVMTPTEGLDTGEMSNSTSSSLKRQRLGSERA 1020

Qy 1021 ASHVAQANLKLLDVSKIFPITAEIABESSSPEVVPVELLCVPSPASOQSDLHTKPLGTDDDFW 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ASHVAQANLKLLDVSKIFPITAEIABESSSPEVVPVELLCVPSPASOGDLHTKPLGTDDDFW 1080

Qy 1081 GPTGPVATEVVDKEKQLYRVHFPVAGSYRWPNTGLCFVVRBEAVTVEIEFCVWDQFLGEIN 1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GPTGPVATEVVDKEKQLYRVHFPVAGSYRWPNTGLCFVVRBEAVTVEIEFCVWDQFLGEIN 1140

Db 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
```

```
Qy   1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200

Qy   1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy   1261  KELELCYRSPGEBQLFSEFYVGHLGSGIRLQVVKDKKDETLVWEALVKPGDLMPATTLIPP  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  KELELCYRSPGEBQLFSEFYVGHLGSGIRLQVVKDKKDETLVWEALVKPGDLMPATTLIPP  1320

Qy   1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLLHGQVLSQBQYERVLAENTRPS  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLLHGQVLSQBQYERVLAENTRPS  1380

Qy   1381  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
```

RESULT 3
ABG78472

ID   ABG78472 standard; Protein; 1429 AA.
XX
AC   ABG78472;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Human; human leucine-rich repeat small intestine I; HLRRSI1; asthma;
KW   Leucine-rich repeat (LRR) domain containing protein #2.
XX
KW   proliferative disorder; gastrointestinal disorder; renal disorder;
KW   neural disorder; reproductive disorder; calcium regulation; apoptosis;
KW   immune system; anaemia; human immune deficiency virus; HIV; cancer;
KW   blood coagulation disorder; autoimmune disorder; allergic reaction;
KW   inflammatory condition; cardiovascular disorder; ischaemia;
KW   neurological disorder; infectious disease; cytokine production;
KW   expressed sequence tag; EST.
XX
OS   Unidentified.
XX
PN   WO20261086-A2.
XX
PD   08-AUG-2002.
XX
PF   20-DEC-2001; 2001WO-US49739.
XX
PR   22-DEC-2000; 2000US-257774P.
XX
PA   (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder J, Ramanathan C, Mintier G;
XX
DR   WPI; 2002-619252/66.
XX
PT   New isolated nucleic acid molecules encoding HLRRSI1 polypeptides, or
PT   their fragments and homologues, useful for preventing, treating and
PT   ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT   or renal disorders -
XX
PS   Disclosure; Page 229-233; 336pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules (I) encoding
CC   human leucine-rich repeat small intestine I (HLRRSI1) polypeptides.
CC   The nucleic acid molecules and polypeptides are useful for preventing,
CC   treating and ameliorating medical conditions, such as proliferative,
CC   gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC   related to aberrant calcium regulation or apoptosis modulation, either
CC   directly or indirectly. They are also useful for treating, preventing
CC   and/or diagnosing diseases, disorders and/or conditions of: immune system
CC   by activating or inhibiting the proliferation, differentiation, or

CC   mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
CC   anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC   virus (HIV) infection, HTLV-BLV infection, blood coagulation disorders,
CC   e.g. arterial thrombosis; asthma or allergic reactions, e.g. Addison's disease,
CC   myasthenia gravis; asthma or allergic reactions; inflammatory
CC   conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC   e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC   ischaemias; aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC   Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC   pneumonia, or viral, bacterial, and fungal infections. The HLRRSI1
CC   polypeptides are useful for modulating cytokine production, antigen
CC   presentation, or other processes such as boosting immune responses.
CC   ABG78454-ABG78474 represent HLRRSI1 amino acid sequences and related
CC   amino acid sequences of the invention.
XX
SQ   Sequence 1429 AA;

```
Qy      1  MAGGAWGRLACYLEFLKKBELKEFQLLLANKAHSRSSSGETPAQPEKTSGGMEVASYLVAQ  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MAGGAWGRLACYLEFLKKBELKEFQLLLANKAHSRSSSGETPAQPEKTSGGMEVASYLVAQ  60

Qy     61  YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSOQPTSTAVLMPW  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSOQPTSTAVLMPW  120

Qy    121  IHELPAGCTQGSERRVLRQLPDTSGRRVWREISASLLYQALPSSSPDHESPSQESPNAPTST  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  IHELPAGCTQGSERRVLRQLPDTSGRRVWREISASLLYQALPSSSPDHESPSQESPNAPTST  180

Qy    181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBKKSEKGRPPWAAVGT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBKKSEKGRPPWAAVGT  240

Qy    241  PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300

Qy    301  PDIYEENRGHLIEIRDLGFGPGLDTQBPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  PDIYEENRGHLIEIRDLGFGPGLDTQBPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR  360

Qy    361  FQHVFYFSCRELAQSKVVSLAELIGKDGFTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  FQHVFYFSCRELAQSKVVSLAELIGKDGFTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420

Qy    421  EPSSELCLHWSQPQPADALLGSLLGKTLPEASFIITARTTALQNLIPSLEQARWVELG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  EPSSELCLHWSQPQPADALLGSLLGKTLPEASFIITARTTALQNLIPSLEQARWVELG  480

Qy    481  FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL  540

Qy    541  TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDLRKHGLDGAI  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDLRKHGLDGAI  600

Qy    601  ISTFLKMGILQEHPIPLSYSFIHLCFQBFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  ISTFLKMGILQEHPIPLSYSFIHLCFQBFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660

Qy    661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWYPSLQLLLQPHSLESLH  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWYPSLQLLLQPHSLESLH  720

Qy    721  CLYEYERNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CLYEYERNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780
```

Query Match   100.0%;   Score 7534;   DB 23;   Length 1429;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1429;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     781 PTMVLFRNWPVTDAYWQLLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 PTMVLFRNWPVTDAYWQLLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy     841 TLRLAGCGLITAEDCKDLAFGLRANQTLITELDLSPNVLTDDAGAKHLCQRLRQPSCKLQRLQ 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TLRLAGCGLITAEDCKDLAFGLRANQTLITELDLSPNVLTDPAGAKHLCQRLRQPSCKLQRLQ 900

Qy     901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ 960

Qy     961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTPEGLDTGEMSNSTSSLKRQRLGSERA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTPEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy    1021 ASHVAQANLKLLDVSKIPPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ASHVAQANLKLLDVSKIPPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy    1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGBIN 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGBIN 1140

Qy    1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200

Qy    1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1260

Qy    1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320

Qy    1321 ARIAVPSPLDAPQLLHPVDQYREQLIARVTSVEVVLIDKLHGQVLSQEQYERVLAENTRPS 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ARIAVPSPLDAPQLLHPVDQYREQLIARVTSVEVVLIDKLHGQVLSQEQYERVLAENTRPS 1380

Qy    1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
```

```
RESULT 4
ABB77916
ID   ABB77916 standard; protein; 1429 AA.
XX
AC   ABB77916;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human leucine-rich repeat domain containing protein KIAA0926.
XX
OS   Homo sapiens.
XX
PN   WO200252011-A2.
XX
PD   04-JUL-2002.
XX
PF   20-DEC-2001; 2001WO-US49740.
XX
PR   22-DEC-2000; 2000US-257773P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder J;  Ramanathan C;  Mintier G;
XX
DR   WPI; 2002-566676/60.
XX
PT   New HLRRBM1 nucleic acids for preventing, treating or ameliorating e.g.
PT   proliferative disorders, immune conditions, a disorder related to
PT   aberrant apoptosis modulation or developmental disorders -
XX
PS   Example 1; Page 362-366; 371pp; English.
XX
CC   The present sequence represents a human leucine-rich repeat domain
CC   containing protein, which was used as a probe to search for leucine-rich
CC   repeat containing protein HLRRBM1. HLRRBM1 polypeptides and
CC   polynucleotides are useful for preventing, treating or ameliorating a
CC   medical condition such as a proliferative disorder, immune condition, or
CC   a disorder related to aberrant apoptosis modulation, either directly or
CC   indirectly, and in modulating signal transduction activity in various
CC   cell, tissue and organisms. They are also useful for treating,
CC   preventing, or diagnosing diseases of haematopoietic cells, autoimmune
CC   disease, graft-versus-host disease, allergic conditions (e.g. asthma),
CC   cardiovascular disorders, and neurological diseases, and for increasing
CC   the organisms' ability to synthesize and/or release pheromones. The
CC   polypeptide may also be used in treating, preventing or ameliorating
CC   rhinitis, or bronchial hyperresponsiveness), reproductive disease,
CC   pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic
CC   haematopoietic platelet disorders (e.g. Bernard-Soulier
CC   syndrome), non-infectious disorders (e.g. innate immunity to bacterial
CC   pathogens, or adaptive immune response), immune and inflammatory
CC   disorders (e.g. Systemic lupus erythematosus), cardiovascular diseases
CC   and cancers. HLRRBM1 nucleic acids may further be used in chromosome
CC   identification or mapping, as a chromosome marker, as molecular weight
CC   markers, as diagnostic probes, in gene therapy, in raising anti-DNA
CC   antibodies, or as antigens for eliciting immune responses.
XX
SQ   Sequence 1429 AA;
```

```
Query Match             100.0%;  Score 7534;  DB 23;  Length 1429;
Best Local Similarity   100.0%;
Matches 1429;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy       1 MAGGAWGRLACYLEPLKKEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MAGGAWGRLACYLEFLKKEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy      61 YGEQRAWDLALHTWEQMGLRSLCAQAEGAGHSPSPPYSPSEPHLGSPSQPTSTAVLMPW 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 YGEQRAWDLALHTWEQMGLRSLCAQAEGAGHSPSPPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy     121 IHELPAGCTQGSERVLRQLPDTSGRRWREISASLIYQALPSSPDHESPSQOESPNAPTST 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 IHELPAGCTQGSERVLRQLPDTSGRRWREISASLIYQALPSSPDHESPSQOESPNAPTST 180

Qy     181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYTEIRERERERSEKGRPPWAAVVGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYTEIRERERERSEKGRPPWAAVVGT 240

Qy     241 PPQAHTSLQPHHHVPWEPSVREBSLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PPQAHTSLQPHHHVPWEPSVREBSLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy     301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVTLQGRAAGIGKSTLARQVKEANGRGQLYGDR 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVTLQGRAAGIGKSTLARQVKEANGRGQLYGDR 360

Qy     361 FQHVFYFSCREFLAQSKVVSLAELIGKDGTATPAPIRQOILSRPERLLFPLDGVDEPGWVLQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 FQHVFYFSCREFLAQSKVVSLAELIGKDGTATPAPIRQOILSRPERLLFPILDGVDEPGWVLQ 420

Qy     421 EPSSELCLHKSQPQPADALLGSLLGKTILPEASFLITARTTALQNILPSLEQARWEVLG 480
```

```
RESULT 5
ABG71631
ID   ABG71631 standard; Protein; 1429 AA.
XX
AC   ABG71631;
```

```
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVELLG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 FSESSRKEYFYRYFTDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQOMRKEKL   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQOMRKEKL   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 TLTSKTTTTLCIHYLAQALCAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTLCIHYLAQALCAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 CLYETRNKIFPLTQVMAHFEEBMGMCVETDMELLLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CLYETRNKIFPLTQVMAHFEEBMGMCVETDMELLLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVRSLCKTLRRPRCLLE  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVRSLCKTLRRPRCLLE  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 TLRLAGCGLTAEDCKDLAFGLRANQCTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TLRLAGCGLTAEDCKDLAFGLRANQCTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 LVSCGLTSDCCQDLASVLSASPSTLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 LVSCGLTSDCCQDLASVLSASPSTLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 TTLSDEMRQELRALIEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTLSDEMRQELRALIEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021 ASHVAQANLKLLDVSKIFPITAEIAEBESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ASHVAQANLKLLDVSKIFPITAEIAEBESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP   1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP   1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVISVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVISVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
```

```
XX
SQ   Sequence   1429 AA;

Query Match             100.0%;  Score 7534;  DB 24;  Length 1429;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1429;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 MAGGAWGRLACYLEFLKKEELKEPQLLIANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAGGAWGRLACYLEFLKKEELKEPQLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy    61 YGEQRAWDLALHTWEQMGLRSLCAQAQEQAGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180

Qy   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERRERKSEKGRPPWAAVVGT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERREKSEKGRPPWAAVVGT 240

Qy   241 PPQAHTSLQPHHHPWPEPSVRESLCSTWPMKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 PPQAHTSLQPHHHPWPEPSVRESLCSTWPMKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy   301 PDYVEENRGHLIEIRDLPGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYSDR 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PDYVEENRGHLIEIRDLPGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYSDR 360

Qy   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASPLITARTTALQNLIPSLEQARWVEVLG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASPLITARTTALQNLIPSLEQARWVEVLG 480

Qy   481 FSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540

Qy   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600

Qy   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy   661 GIHGLFGASTTRFLLGILSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GIHGLFGASTTRFLLGILSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy   721 CLYPTRNKTPLTQVMAHPEMGMCVETDMELLVCTPCIKPSRHVKKLQLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CLYPTRNKTPLTQVMAHPEMGMCVETDMELLVCTPCIKPSRHVKKLQLIEGRQHRSTWS 780

Qy   781 PTMVVLFRNVPVTDAYWQLLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 PTMVVLFRNVPVTDAYWQLLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy   841 TLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTDAGAKHLCQRLROPSCKLQRLQ 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTDAGAKHLCQRLROPSCKLQRLQ 900

Qy   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy   961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy  1021 ASHVAQANLKLLDVSKIFIIAEIAEESSEVVPYELLCVPSPASQGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ASHVAQANLKLLDVSKIFIIAEIAEESSEVVPYELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy  1081 GPTGPVATEVVDKERQNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GPTGPVATEVVDKKRQNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140

Qy  1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200

Qy  1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYRYVHPEEVTFHLYLIIPSDCSIR 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYRYVHPEEVTFHLYLIIPSDCSIR 1260

Qy  1261 KELELCYRSPGEDQLFSEFYVGHLGSSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 KELELCYRSPGEDQLFSEFYVGHLGSSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP 1320

Qy  1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQBQYERVLAENTRPS 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQBQYERVLAENTRPS 1380

Qy  1381 QMRKLFSLSQSWDRKCCKDGLYQALKETHPHLIMELWEKGSKKGLLLPLSS 1429
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 QMRKLFSLSQSWDRKCCKDGLYQALKETHPHLIMELWEKGSKKGLLLPLSS 1429
```

```
RESULT 6
ABG71633
ID   ABG71633 standard; Protein; 1429 AA.
XX
AC   ABG71633;
XX
DT   10-JAN-2003  (first entry)
XX
DE   Human caspase recruitment domain-7 (CARD-7).
XX
KW   Human; caspase activity; caspase recruitment domain-7; CARD-7;
KW   caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-lbeta;
KW   pseudoICE; ICEBERG; cell growth; cell death; inflammation;
KW   apoptosis; caspase activation; cancer; follicular lymphoma;
KW   leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
KW   autoimmune disease; systemic lupus erythematosus; reactive arthritis;
KW   human immunodeficiency virus infection; HIV infection; ALS;
KW   Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW   myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
KW   insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
KW   graft rejection; allergic rhinitis; food allergy; conjunctivitis;
KW   glomerular nephritis; cytostatic; virucide; immunosuppressive;
KW   dermatological; nephrotropic; neuroprotective; cardiant.
OS   Homo sapiens.
XX
PN   US2002128198-A1.
XX
PD   12-SEP-2002.
XX
PF   27-NOV-2001; 2001US-0996617.
XX
PR   28-JUN-1999;  99US-0340620.
PR   27-OCT-1999;  99US-0428252.
PR   15-AUG-2001; 2001US-0931071.
XX
PA   (BERT/) BERTIN J.
XX
PI   Bertin J;
XX
DR   WPI; 2003-028967/02.
DR   N-PSDB; ABS56030.
```

```
XX
PT   Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT   CARD-7 and CARD-5 in presence of test compound, measuring their
PT   binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PT   altered -
XX
PS   Disclosure; Fig 1; 43pp; English.
XX
CC   The present invention relates to methods of identifying compounds
CC   that regulate caspase activity using caspase recruitment domain-7
CC   (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC   method for identifying a compound that modulates the interaction
CC   between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC   converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
CC   CARD-8 molecules are useful as modulating agents in regulating a
CC   variety of cellular processes including cell growth, cell death, and
CC   inflammation. The methods of the invention are useful for identifying
CC   compounds that have the ability to increase/decrease apoptosis, or
CC   comprise the ability to induce caspase activation. The methods are
CC   useful for treating a disorder associated with inappropriate apoptosis
CC   or inappropriate inflammation. The methods are useful for treating
CC   disorders associated with an undesirably low rate of apoptosis such
CC   as cancer (preferably follicular lymphoma, chronic myelogenous
CC   leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
CC   infections, autoimmune diseases caused by low levels of apoptosis
CC   (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC   and arthritis). The methods are also useful for treating disorders with
CC   undesirably high rates of apoptosis such as human immunodeficiency
CC   virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC   amylotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC   muscular atrophy, various forms of cerebellar degeneration, anaemia
CC   associated with chronic disease, aplastic anaemia, chronic neutropaenia,
CC   myelodysplastic syndromes, myocardial infarction, stroke, and
CC   various inflammatory disorders (e.g. Crohn's disease, reactive
CC   arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC   psoriasis, graft rejection, allergic rhinitis, food allergies,
CC   conjunctivitis, glomerular nephritis, etc). The present sequence
CC   represents human CARD-7.
XX
SQ   Sequence 1429 AA;

Query Match             100.0%;  Score 7534;  DB 24;  Length 1429;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  1429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy         61 YGEQORAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YGEQORAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy        121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDPHESPSQESNNAPTST 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDPHESPSQESNAPTST 180

Qy        181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBRKSEKGRPPWAAVVGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBRKSEKGRPPWAAVVGT 240

Qy        241 PPQAHTSLQPHHPWPEPSVREBSLCCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPQAHTSLQPHHPWPEPSVREBSLCCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy        301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy        421 EPSSELCLHWSQPQPADALLGSLLGKTIILPEASFLITARTTALQNLIPSLEQARWEVIG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EPSSELCLHWSQPQPADALLGSLLGKTIILPEASFLITARTTALQNLIPSLEQARWEVIG 480

Qy        481 FSBSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FSBSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540

Qy        541 TLITSKITTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TLITSKITTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy        601 ISTFLKMGILQEHPIPLSYSPIHLCFQEEFFAAMSXYLEDEKGRGKHISNCIIDLEKTLEAY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ISTFLKMGILQEHPIPLSYSPIHLCFQEEFFAAMSXYLEDEKGRGKHISNCIIDLEKTLEAY 660

Qy        661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLBSLH 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLBSLH 720

Qy        721 CLYETRNKTFLITQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CLYETRNKTFLITQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780

Qy        781 PTMVVLFRWPVTDAYWQILLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PTMVVLFRWPVTDAYWQILLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840

Qy        841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLRIRLGLDQ 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLRIRLGLDQ 960

Qy        961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTTEGLDTGEMSNSTSSLKRQRLGSERA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy       1021 ASHVAQANLKLLDVSKIFPITABIAEBSSPEVVPVELLCVFSPASQGDLHTKPLGTDDFW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ASHVAQANLKLLDVSKIFPITABIAEBSSPEVVPVELLCVFSPASQGDLHTKPLGTDDFW 1080

Qy       1081 GPTGPVATEVVDKEKNLYRVHPPVAGSYRWPNTGICFVMREAVTVEIBFCVWDQFLGEIN 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GPTGPVATEVVDKEKNLYRVHPPVAGSYRWPNTGICFVMREAVTVEIBFCVWDQFLGEIN 1140

Qy       1141 PQHSWMVAGPLLDIKAEPGAVEEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 PQHSWMVAGPLLDIKAEPGAVEEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200

Qy       1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIR 1260

Qy       1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVRDKDDETLVWEALVKPGDLMPATTLIPP 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVRDKDDETLVWEALVKPGDLMPATTLIPP 1320

Qy       1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy       1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
```

RESULT 7
AAY72669
ID   AAY72669 standard; Protein; 1473 AA.
XX

us-09-996-617-2.rag

```
AC   AAY72669;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NB-ARC and CARD containing protein (NAC) beta isoform.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FH
FT   Domain          329..547
FT                   /label= NB domain
FT                   /note= "Nucleotide binding domain, also designated as
FT                   NB_ARC domain"
FT   Domain          329..341
FT                   /label= walker_A
FT                   /note= "Also designated as P-loop"
FT   Domain          406..414
FT                   /label= walker_B
FT   Region          809..833
FT                   /label= Leucine_rich_repeat_region
FT   Region          838..862
FT                   /label= Leucine_rich_repeat_region
FT   Region          865..890
FT                   /label= Leucine_rich_repeat_region
FT   Region          895..919
FT                   /label= Leucine_rich_repeat_region
FT   Region          923..947
FT                   /label= Leucine_rich_repeat_region
FT   Region          957..987
FT                   /label= Leucine_rich_repeat_region
FT   Region          1261..1306
FT                   /note= "This 45 amino acid segment is not found in
FT                   NAC gamma isoform (AAY72671) due to alternative
FT                   mRNA splicing"
FT   Region          1261..1306
FT                   /note= "This 31 amino acid segment is not found in
FT                   NAC gamma isoform (AAY72670) and NAC delta isoform
FT                   (AAY72671) due to alternative mRNA splicing"
FT   Domain          1079..1364
FT                   /label= "TIM-Barrel-like domain"
FT   Domain          1128..1473
FT                   /label= CARD-L
FT   Domain          1128..1261
FT                   /note= "Caspase-associated recruitment domain"
FT   Domain          1128..1261
FT                   /label= CARD-S
FT   Domain          1306..1473
FT                   /note= "Caspase-associated recruitment domain"
FT   Domain          1306..1473
FT                   /label= CARD-S
FT   Domain          1373..1473
FT                   /note= "Caspase-associated recruitment domain"
FT   Domain          1373..1473
FT                   /label= CARD
FT                   /note= "Caspase-associated recruitment domain"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
PI   Reed JC;
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
```

```
XX
CC   The present sequence is a human NB-ARC and CARD containing protein
CC   (NAC) beta isoform. NAC beta isoform represents the NAC splice
CC   variant in which both the splice regions are present in the translated
CC   polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC   (also referred as NB-ARC domain), a caspase-associated recruitment
CC   domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
CC   aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC   NAC proteins are used for screening modulators that modulates apoptosis,
CC   cytokine production, cytokine receptor signalling and other cellular
CC   processes. NAC can act as an immunogen for the production of polyclonal
CC   and monoclonal antibodies. It can also be used to diagnose and treat
CC   inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC   pathologies such as adenocarcinomas and leukaemias.
CC   Note: This sequence is stated as being the same as that shown as
CC   SEQ ID NO:2 (AAY72711) in figure 1A of the specification. However the
CC   sequences differ at several positions.
XX
SQ   Sequence   1473 AA;
```

```
XX
Query Match      99.6%;    Score 7502;    DB 22;    Length 1473;
Best Local Similarity    97.0%;    Pred. No.0;
Matches 1429;    Conservative    0;    Mismatches    0;    Indels   44;    Gaps    1;

Qy     1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSGGETPAQPEKTSGMEVASYLIVAQ 60
         |||||||||||||||||||||||||||||||||||||||||
Db     1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLIVAQ 60

Qy    61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSFPYSPSEPHLGSPSQPTSTAVLMPW 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy   121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDHESSPSQGBSPNAPTST 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDHESSPSQGBSPNAPTST 180

Qy   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIREREREKSEKGRPPWAAVGT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIREREREKSEKGRPPWAAVVGT 240

Qy   241 PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDFLVKRSW 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDFLVKRSW 300

Qy   301 PDYEENRGHLIEIRDLFGPGLDTQBPRIVILLQGPAAGIGKSTLARQVKEAWGRGQLYGDR 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PDYEENRGHLIEIRDLFGPGLDTQBPRIVILLQGPAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPBASFLITARTTALQNLIPSLEQARWEVLG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPBASFLITARTTALQNLIPSLEQARWEVLG 480

Qy   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy   541 TLTSKYTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKYTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy   601 ISTFLKMGILQEHFIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHFIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy   661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
```

RESULT 8
AAY72711

```
Db    661  GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH   720
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVXKLQLIEGRQHRSTWS   780
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVXKLQLIEGRQHRSTWS   780

Qy    781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLH   840
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLH   840

Qy    841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ   900
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ   900

Qy    901  LVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRIPACKLIRLGLDQ   960
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  LVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRIPACKLIRLGLDQ   960

Qy    961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA  1020
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA  1020

Qy   1021  ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVEBLLCVPSPASQGDLHTKPLGTDDDFW  1080
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVEBLLCVPSPASQGDLHTKPLGTDDDFW  1080

Qy   1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN  1140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN  1140

Qy   1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV  1200

Qy   1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy   1261  ------------------------------KELELCYRSPGEBDQLF  1276
                                               ||||||||||||||||
Db   1261  KAIDDLEMKFQFVRIHKPPPLTPLVMGCRYTVSGSGSGSGMLEILPKELELCYRSPGEBDQLF  1320

Qy   1277  SEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1336
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  SEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1380

Qy   1337  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKC  1396
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKC  1440

Qy   1397  KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
                |||||||||||||||||||||||||||||||||
Db   1441  KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1473
```

```
CC   aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC   NAC proteins are used for screening modulators that modulates apoptosis,
CC   cytokine production, cytokine receptor signalling and other cellular
CC   processes. NAC can act as an immunogen for the production of polyclonal
CC   and monoclonal antibodies. It can also be used to diagnose and treat
CC   pathologies such as sepsis, fibrosis and arthritis and cancer
CC   inflammatory disorders such as adenocarcinomas and leukaemias.
CC   Note: This sequence is stated as being the same as that shown as
CC   SEQ ID NO:2 (See AAY72669) in page 133-137 of the specification. However
CC   the sequences differ at several positions.
XX
SQ   Sequence    1473 AA;

Query Match             99.6%;  Score 7502;  DB 22;  Length 1473;
Best Local Similarity   97.0%;  Pred. No. 0;  Mismatches 0;
Matches 1429;  Conservative  0;  Mismatches  0;  Indels  44;  Gaps  1;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPQPTSTAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASILYQALPSSPDHESPSQESPNAPTST 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASILYQALPSSPDHESPSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERERKSEKGRPPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERERKSEKGRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHHWEPSVRESLGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVXRSW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHHWEPSVRESLGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVXRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLIHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLIHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy    541 TLISKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLISKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGKHSNCIIDLEKTLEAY 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLIYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLIYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781 PTMVLFRWVPTDAYWQILFSVLKYTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMVLFRWVPTDAYWQILFSVLKYTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
```

```
Qy    841 TLRLAGCGLITAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLITAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLTSDCCQDLASVLSASPSLKEIDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLTSDCCQDLASVLSASPSLKEIDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy   1021 ASHVAQANLKLLDVSKIFPPIAEIAEESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ASHVAQANLKLLDVSKIFPPIABIAEESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy   1081 GPTGPVATEVVDKEQNLYRVHPPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GPTGPVATEVVDKEQNLYRVHPPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140

Qy   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLPQMAHFKEEGMLLEKPARV 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLPQMAHFKEEGMLLEKPARV 1200

Qy   1201 ELHHIVLENPSFSPLGCVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIIPSDCSIR 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ELHHIVLENPSFSPLGCVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIIPSDCSIR 1260

Qy   1261 K------------------------------------------ELELCYRSPGEDOLF 1276
          |                                          |||||||||||||
Db   1261 KAIDDLEMKFQFVRIHKPPLTPLYMGCRYTVSGSGSRDAGNTPQELELCYRSPGEDOLF 1320

Qy   1277 SEFYVGHLGSGIRLQVRDKDDETLWEEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 SEFYVGHLGSGIRLQVRDKDDETLWEEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1380

Qy   1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQVRKLFSLSQSWDRKC 1396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQVRKLFSLSQSWDRKC 1440

Qy   1397 KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
          ||||||||||||||||||||||||||||||||
Db   1441 KDGLIYQALKETHPHLIMELWEKGSKKGLLPLSS 1473
```

RESULT 9
ABG78455

```
ID   ABG78455 standard; Protein; 1429 AA.
XX
AC   ABG78455;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human caspase recruitment protein 7 protein.
XX
KW   Human; human leucine-rich repeat small intestine I; HLRRSI1; asthma;
KW   proliferative disorder; Gastrointestinal disorder; renal disorder;
KW   neural disorder; reproductive disorder; calcium regulation; apoptosis;
KW   immune system; anaemia; human immune deficiency virus; HIV; cancer;
KW   blood coagulation disorder; autoimmune disorder; allergic reaction;
KW   inflammatory condition; cardiovascular disorder; ischaemia;
KW   neurological disorder; infectious disease; cytokine production;
KW   expressed sequence tag; EST.
OS   Homo sapiens.
XX
PN   WO20261086-A2.
XX
PD   08-AUG-2002.
XX
PF   20-DEC-2001; 2001WO-US497739.
XX
PR   22-DEC-2000; 2000US-257774P.
```

```
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder J, Ramanathan C, Mintier G;
XX
DR   WPI; 2002-619252/66.
XX
PT   New isolated nucleic acid molecules encoding H1RRSI1 polypeptides, or
PT   their fragments and homologues, useful for preventing, treating and
PT   ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT   or renal disorders
XX
PS   Disclosure; Figure 2; 336pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules (I) encoding
CC   human leucine-rich repeat small intestine I (HLRRSI1) polypeptides.
CC   The nucleic acid molecules and polypeptides are useful for preventing,
CC   treating and ameliorating medical conditions, such as proliferative,
CC   gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC   related to aberrant calcium regulation or apoptosis modulation, either
CC   directly or indirectly. They are also useful for treating, preventing
CC   and/or diagnosing diseases, disorders and/or conditions of: immune system
CC   by activating or inhibiting the proliferation, differentiation, or
CC   mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
CC   anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC   virus (HIV) infection, HTLV-BLV infection; blood coagulation disorders,
CC   e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC   myasthenia gravis; asthma or allergic reactions; inflammatory
CC   conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC   e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC   ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC   Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC   pneumonia, or viral, bacterial, and fungal infections. The HLRRSI1
CC   polypeptides are useful for modulating cytokine production, antigen
CC   presentation, or other processes such as boosting immune responses.
CC   ABG78454-ABG78474 represent HLRRSI1 amino acid sequences and related
CC   amino acid sequences of the invention.
XX
SQ   Sequence 1429 AA;

Query Match             99.4%;  Score 7488;  DB 23;  Length 1429;
Best Local Similarity   99.4%;
Matches 1420;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 MAGGAWGRLACYLFPLKKEELKEFQLLLANKAHSRSSGGETPAQPEKTSGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLFPLKKEELKEFQLLLANKAHSRSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
          ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db    121 IHELPAGCTQGGSERRVLRQLPDTSGRRWREISASHLYQALPSSPDHESPSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERREKSEKGRPPWAAVVGT 240
          ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDBTSGIYYTEIRERERREKSEKGRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          |||||:||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db    241 PPQAHSSLQPHHHPWEPSVRESLCSTWPWKNEEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLFWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLFWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy    481 FSEESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSEESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy    541 TLTSKTTTTLCLHYLAQALCQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLTSKTTTTLCLHYLAQALCQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHHIPLSYSFPIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHHIPLSYSFPIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLYETRNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLLQLIEGRQHRSTWS 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYETRNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLLQLIEGRQHRSTWS 780

Qy    781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PSMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTMDAGAKHLCQRLRQPSCKLQRLQ 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTMDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGIRHPACKLIRLGLDQ 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGIRHPACKLIRLGLDQ 960

Qy    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTFTEGLDTGEMSNSTSSLKRQRLGSERA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTFTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy   1021 ASHVAQANLKLLDVSKIFPIAEIRAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ASHVAQANLKLLDVSKIFPIAEIRAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy   1081 GPTGGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREATVTVEIEFCVWDQFLGEIN 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GPTGGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREATVTVEIEFCVWDQFLGEIN 1140

Qy   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
          ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQVAHFKEEGMLLEKPARV 1200

Qy   1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYRHVHPEEVTFHLYLIPSDCSIR 1260
          ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db   1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYRHLHPEEVTFHLYLIPSDCSIR 1260

Qy   1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVRDKDKDETLVWEALVKPGDLMPATTLIPP 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVRDKDKDETLVWEALVKPGDLMPATTLIPP 1320

Qy   1321 ARIAVPSPLDAPQLIHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
          |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ACIAVPSPLDAPQLIHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy   1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429

RESULT 10
AAO17855
ID   AAO17855 standard; Protein; 1429 AA.
XX
AC   AAO17855;
```

```
XX
DT   20-AUG-2002 (first entry)
XX
DE   Pyrin domain containing protein NALP1-hs.
XX
KW   Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW   antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW   neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW   nephrotropic; nootropic; osteopathic; intracellular signal transduction;
KW   inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW   arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW   osteoarthritis; glomerulonephritis.
XX
OS   Unidentified.
XX
PN   WO200240668-A2.
XX
PD   23-MAY-2002.
XX
PF   30-OCT-2001; 2001WO-EP12545.
XX
PR   15-NOV-2000; 2000DE-1056687.
PR   30-NOV-2000; 2000DE-1059595.
XX
PA   (APOT-) APOTECH RES & DEV LTD.
XX
PI   Tschopp J, Martinon F;
XX
DR   WPI; 2002-427093/45.
DR   N-PSDB; AAL47127.
XX
PT   New DNA encoding protein with pyrin domain, useful for treating
PT   diseases involving impaired signal transduction, particularly
PT   inflammation, also proteins and antibodies -
XX
PS   Claim 5; Fig 1; 116pp; German.
XX
CC   The present invention relates the DNA and their encoded proteins, where
CC   the proteins contain at least one PYD (pyrin) domain. These can be used
CC   to treat diseases associated with impaired intracellular signal
CC   transduction, particularly inflammation such as psoriasis,
CC   arteriosclerosis, bacterial or viral infections (particularly meningitis
CC   and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC   sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC   and Parkinson's diseases. The present sequence is a protein of the
CC   invention.
XX
SQ   Sequence   1429 AA;

Query Match             99.4%;  Score 7488;  DB 23;  Length 1429;
Best Local Similarity   99.4%;
Matches 1420;  Conservative  5;  Mismatches  4;  Indels  0;  Gaps  0;

Qy     1 MAGGAWGRLACYLEFLKEELKEFQLLLANKAHSRSSSSGETPAQPEKTSGMEVASYLVAQ 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAGGAWGRLACYLEFLKEELKEFQLLLANKAHSRSSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy    61 YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESFSPSQESPNAPTST 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASHLYQALPSSPDHESPSQESSNAPTST 180

Qy   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy   241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDENVQKFFQLLLLQRPHPRSQDPLVKRSW 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 PPQAHSSLQPHHHPWEPSVRESLCSTWPWKNEDRDNQKFFQLLLLQRPHPRSQDPLVKRSW 300

Qy   301 PDYVEENRGHLIEIRDLFGPSLDTQEPRIVILQGGAAGIGKSTLARQVKEANGRGQLYGDR 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PDYVEENRGHLIEIRDLFGPSLDTQEPRIVILQGGAAGIGKSTLARQVKEANGRGQLYGDR 360

Qy   361 FQHVFYFSCRELAQSKVSLAEBLIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FQHVFYFSCRELAQSKVSLAEBLIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWEVLG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWEVLG 480

Qy   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKKEKL 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKKEKL 540

Qy   541 TLTSKTTTTLCLHYLAQALCQAOPLGPQOLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTLCLHYLAQALCQAOPLGPQOLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600

Qy   601 ISTFLKMGILQEHPIPLSYSFIHLCPQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHPIPLSYSFIHLCPQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy   661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy   721 CLYETRNKTFLITQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CLYETRNKTFLITQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy   781 PTMVVLFRWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
         |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 PSMVVLFRWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy   841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLMDAGAKHLCQRLRQPSCKLQRLQ 900

Qy   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLRIRLGLDQ 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLRIRLGLDQ 960

Qy   961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLRQRLGSERA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPIEGLDTGEMSNSTSSLRQRLGSERA 1020

Qy  1021 ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVRELLCVPSPASQCDLHTKPLGTDDDFW 1080

Qy  1081 GPTGPVATEVVDKEKNLYRVHFPPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVVREAVTVEIEFCVWDQFLGEIN 1140

Qy  1141 PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSLFQVAHFKEEGMLLEKPARV 1200

Qy  1201 ELHHIVLENFSFSPLGVLLRMIHNALRFIPVTSVLLIYRHVHPEEVTFHLYLIPSDCSIR 1260
         |||||||||||||||||||||||||||||||:||:|||||||||||||||||||||||||
Db  1201 ELHHIVLENFSFSPLGVLLKMIHNALRFIPVTSVLLYRHLHPEEVTFHLYLIPSDCSIR 1260

Qy  1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320

Qy  1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
         ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db  1321 ACIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy  1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
```

```
Db          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            1381 QMRKLFSLSQSWWDRKCKDGLYQALKETHPHLIMELWEKGSKGLLPLSS 1429

RESULT 11

ID    AAE06758 standard; Protein; 1473 AA.
XX
AC    AAE06758;
XX
DT    16-OCT-2001 (first entry)
XX
DE    Human G-protein coupled receptor-8 (GCREC-8) protein.
XX
KW    Human; G-protein coupled receptor-8; GCREC-8; cytostatic; hepatotropic;
KW    virucide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective;
KW    nootropic; cerebroprotective; hypotensive; tranquilliser; vulnerary;
KW    ophthalmological; cell proliferative disorder; actinic keratosis;
KW    anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
KW    psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease;
KW    Huntington's disease; Parkinson's disease; cardiovascular disorder;
KW    epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia;
KW    anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder;
KW    Addison's disease; Crohn's disease; acquired immune deficiency syndrome;
KW    AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity;
KW    osteoporosis; transgenic animal; gene therapy.
XX
OS    Homo sapiens.
XX
FH    Key             Location/Qualifiers
FH
FT    Domain          1216..1237
FT                    /label= Transmembrane_domain
FT    Binding-site    334..341
FT                    /label= ATP/GTP-binding_site
FT                    /note= "P-loop"
XX
PN    WO200157085-A2.
XX
PD    09-AUG-2001.
XX
PF    01-FEB-2001; 2001WO-US03455.
XX
PR    02-FEB-2000; 2000US-0180093.
PR    11-FEB-2000; 2000US-0182045.
XX
PA    (INCY-) INCYTE GENOMICS INC.
XX
PI    Baughn MR, Au-Young J, Yue H;
XX
DR    WPI; 2001-488869/53.
DR    N-PSDB; AAD12951.
XX
PT    Novel isolated human G-protein coupled receptor-8 (GCREC-8)
PT    protein. The present invention relates to GCREC protein and nucleic
PT    acids encoding them. GCREC protein, its agonist or antagonist are useful
PT    for treating diseases or conditions associated with decreased expression
PT    or overexpression of functional GCREC in a patient, where the disorder
PT    is selected from cell proliferative disorders such as actinic keratosis,
PT    arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
PT    cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
PT    disease, Huntington's disease, Parkinson's disease, cardiovascular
PT    disorders such as hypertension, vasculitis, varicose veins, gastro-
PT    intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
PT    pancreatitis, autoimmune/inflammatory disorders such as acquired
PT    immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
PT    uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
```

```
Qy    Sequence    1473 AA;
SQ

            Query Match        99.3%;  Score 7484;  DB 22;  Length 1473;
            Best Local Similarity   96.8%;  Pred. No. 0;
            Matches 1426;  Conservative   1;  Mismatches   2;  Indels   44;  Gaps   1;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSPEHLGSPSSQPTSTAVLMPW 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSPEHLGSPSSQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
          ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASHVYQALPSSPDHESPSQESPNAPTST 180

Qy    181 AVLGSWGSPPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHHPWEPSVRESLCSTWPKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHHPWEPSVRESLCSTWPKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301 PDVVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLLARQVKEAWGRGQLYGDR 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDVVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQLLSRPERLLFILDGVDEPGWVLQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQLLSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWILACTCLMQQMKREKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWILACTCLMQQMKREKL 540

Qy    541 TLTSKTTTTLCLHYLAQALQAQPLGPQOLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLTSKTTTTLCLHYLAQALQAQPLGPQOLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLYEFTRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYEFTRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781 PTMVVLFRWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMVVLFRWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840

Qy    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKILCQRLRQPSCKLQRLQ 900
```

```
CC    infections, trauma and metabolic disorders such as diabetes, obesity,
CC    osteoporosis, trauma and metabolic disorders and their cDNAs are used to assess the
CC    effects of exogenous compounds on the expression of GCREC sequences.
CC    GCREC cDNA is useful to create knock in humanised animals (pigs) or
CC    transgenic animals (mice or rats) to model human disease, for
CC    therapeutic or diagnostic purposes, for somatic or germline gene
CC    therapy, to generate hybridisation probes useful in mapping the
CC    naturally occurring genomic sequence, and in molecular biological
CC    techniques.
```

```
PS    Claim 1; Page 114-117; 138pp; English.
```

```
Db   841  TLRLAGCLTAEDCKDLAEGIRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960

Qy   961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRORLGSSERA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRORLGSSERA 1020

Qy  1021  ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080

Qy  1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVRREAVTVEIEFCVWDQFLGEIN  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVRREAVTVEIEFCVWDQFLGEIN  1140

Qy  1141  PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSLFQWAHFKEEGMLLEKPARV 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSLFQWAHFKEEGMLLEKPARV 1200

Qy  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy  1261  ------------------------------KELELCYRSPGEDQLF               1276
                                        ||||||||||||||
Db  1261  KAIDDLEMKFQVRIHKPPLTPLYMGCRYTVSGSGSGMLEILPKELELCYRSPGEDQLF    1320

Qy  1277  SEFVVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1336
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  SESYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1380

Qy  1337  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEBQYERVLAENTRPSQMRKLFSLSQSWDRKC 1396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEBQYERVLAENTRPSQMRKLFSLSQSWDRKC 1440

Qy  1397  KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS   1429
          |||||||||||||||||||||||||||||||||
Db  1441  KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS   1473
```

RESULT 12

AAY72670

ID   AAY72670 standard; Protein; 1397 AA.
XX
AC   AAY72670;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NB-ARC and CARD containing protein (NAC) gamma isoform.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   Region   1..956
FT   /note="Corresponds to 1-956 residues of human NAC
FT   beta isoform (AAY72669)"
FT   Region   957..1229
FT   /note="Corresponds to 988-1260 residues of human NAC
FT   beta isoform (AAY72669)"
FT   Region   1230..1397
FT   /note="Corresponds to 1306-1473 residues of human NAC
FT   beta isoform (AAY72669)"
XX

PN   WO20011617O-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
AC   AAY72670;
PF   01-SEP-1999; 99US-0382221.
XX
PR   (BURN-) BURNHAM INST.
XX
PA   Reed JC;
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   N-PSDB; AAD02761.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
PS   Claim 15; Page 143-148; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein
CC   (NAC) gamma isoform. NAC gamma isoform represents the NAC splice
CC   variant in which both the splice regions are absent in the translated
CC   polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC   (also referred as NB-ARC domain), a caspase-associated recruitment
CC   domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
CC   aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC   NAC proteins are used for screening modulators that modulates apoptosis,
CC   cytokine production, cytokine receptor signalling and other cellular
CC   processes. NAC can act as an immunogen for the production of polyclonal
CC   and monoclonal antibodies. It can also be used to diagnose and treat
CC   inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC   pathologies such as adenocarcinomas and leukaemias.
XX
SQ   Sequence 1397 AA;

Query Match   97.5%; Score 7342; DB 22; Length 1397;
Best Local Similarity   97.8%; Pred. No. 0; Mismatches 0;
Matches 1397; Conservative 0; Mismatches 0; Indels 32; Gaps 2;

```
Qy     1  MAGGAWGRLACYLEFLKCEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ    60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MAGGAWGRLACYLEFLKCEELKEFQLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ    60

Qy    61  YGEQRAWDLALHTWEQMGLRSLCAQQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  YGEQRAWDLALHTWEQMGLRSLCAQQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW   120

Qy   121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180

Qy   181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERERKSEKGRPPWAAVVGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERERKSEKGRPPWAAVVGT 240

Qy   241  PPQAHTSLQPHHHPWEPSVRESLCSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  PPQAHTSLQPHHHPWEPSVRESLCSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300

Qy   301  PDYVEENRGHLIEIRDLFPGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  PDYVEENRGHLIEIRDLFPGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR    360

Qy   361  FQHVFYFPSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  FQHVFYFPSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420

Qy   421  EPSSEICLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWEVLG    480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  EPSSEICLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWEVLG    480
```

```
Qy    481  FSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  FSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL  540

Qy    541  TLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI   600

Qy    601  ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY   660

Qy    661  GIHGLFGASTTRFLIGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GIHGLFGASTTRFLIGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH   720

Qy    721  CLYETRNKIFLTQVMAHFEBMGMCVETDMELLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS  780
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    721  CLYETRNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS  780

Qy    781  PTMVVLFRWVPVTDAYWQILLFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE  840
           ||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||
Db    781  PTMVVLFRNVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE   840

Qy    841  TLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTDDAGAKHLCQRLRQPSCKLQRLQ  900
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    841  TLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ   900

Qy    901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRL---   956

Qy    961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTBGLDTGEMSNSTSSLKRQRLGSERA  1020
                             |||||||||||||||||||||||||||||||||||
Db    957  ------------------KPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA          989

Qy   1021  ASHVAQANLKLLDVSKIFPILAEIAEBSSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    990  ASHVAQANLKLLDVSKIFPILAEIAEBSSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1049

Qy   1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1109

Qy   1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQCQGHVDTSLFQMAHFKEEGMLLEKPARV 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQCQGHVDTSLFQMAHFKEEGMLLEKPARV 1169

Qy   1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1229

Qy   1261  KELELCYRSPGEDQLFSEFVYGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP   1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230  -ELELCYRSPGEDQLFSEFVYGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP   1288

Qy   1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1289  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1348

Qy   1381  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS             1429
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1349  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS             1397
```

Query Match   97.2%; Score 7325.5; DB 22; Length 1442;
Best Local Similarity   94.9%;
Matches 1398; Conservative   0; Mismatches   0; Indels   75; Gaps 2;

```
Qy    1  MAGGAWGRLACYLFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ    60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MAGGAWGRLACYLFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ    60

Qy   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW  120

Qy  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESPSQESPNAPTST 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESPSQESPNAPTST 180

Qy  181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERREKSEKGRPPWAAVVGT 240
```

```
Qy    181  AVLGSWGGSPPQPSLAPRREQEAPGTQWPLDETSGIYTEIRERERKSEKGRPPWAAAVVGT 240

Qy    241  PPQAHTSLQPHHPWPEPSVRESLCSTWPKKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  PPQAHTSLQPHHPWPEPSVRESLCSTWPKKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301  PDVVEBNRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  PDVVEBNRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361  FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWLQ 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWLQ 420

Qy    421  EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy    481  FSESSRKEYFYRYLFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKREKL 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  FSESSRKEYFYRYLFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKREKL 540

Qy    541  TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601  ISTFLKQGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  ISTFLKQGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQHSLESLH 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQHSLESLH 720

Qy    721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy    841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLITDAGAKHLCQRLRQPSCKLQRLQ 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLITDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRL---- 956

Qy    961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLLDTGEMSNSTSSLKRQRLGSERA 1020
           ||||||||||||||||||||||||
Db    957  --------------------------KPSVMTPTEGLLDTGEMSNSTSSLKRQRLGSERA 989

Qy    1021 ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990  ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1049

Qy    1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1050 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1109

Qy    1141 PQHSWMVAGPLLDLIKAEPGAVEAVHLPHFVALQGGHVDDTSLFQMAHFKEEGMLLEKPARV 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1110 PQHSWMVAGPLLDLIKAEPGAVEAVHLPHFVALQGGHVDDTSLFQMAHFKEEGMLLEKPARV 1169

Qy    1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1170 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1229

Qy    1261 ------------------------------------------------KELELCYRSPGEDQLF 1276
                                                            |||||||||||
Db    1230 KAIDDLEMKQFVRIHKPPPLPTLYMGCRYTVSGSGSGGMLEILPKELELCYRSPGEDQLF 1289

Qy    1277 SEFVVGHLGSSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1290 SEFVVGHLGSSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1349

Qy    1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWMDRKC 1396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1350 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWMDRKC 1409

Qy    1397 KDGLYQALKETHPHLIMELWEKGSSKCGLLPLSS 1429
           ||||||||||||||||||||||||||||||||||
Db    1410 KDGLYQALKETHPHLIMELWEKGSSKKGLLPLSS 1442
```

RESULT 14
ID   AAY72673   standard; Protein; 1454 AA.
XX
AC   AAY72673;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NAC beta isoform-CARD-X1 chimeric protein.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW   chimeric protein.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT
FT   Region   1..1078
FT   /note= "Corresponds to 1-1078 residues of human NAC
FT   beta isoform (AAY72669)"
FT   Region   1079..1454
FT   /note= "Corresponds to 56-431 residues of human CARD-X
FT   protein (AAY72672)"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   N-PSDB; AAD02764.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
PS   Disclosure; Page 168-172; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC   beta isoform-CARD-X1 chimeric protein. NAC protein comprises a
CC   nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC   caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC   domain. CARD-X protein comprises a caspase-associated recruitment domain
CC   (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC   proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC   for screening modulators that modulates apoptosis, cytokine production.
CC   cytokine receptor signalling and other cellular processes. They can act
CC   as an immunogen for the production of polyclonal and monoclonal
CC   antibodies. They can also be used to diagnose and treat inflammatory

```
CC   disorders such as sepsis, fibrosis and arthritis and cancer pathologies
CC   such as adenocarcinomas and leukaemias.
XX
SQ   Sequence  1454 AA;

Query Match     84.2%;  Score 6344;  DB 22;  Length 1454;
Best Local Similarity  84.2%;  Pred. No. 0;
Matches 1229;  Conservative  53;  Mismatches  122;  Indels  56;  Gaps  7;

Qy      1 MAGGAWGRIACYLEFLKEEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYILVAQ 60
          |||||||.||||||||.|||||||||||||||||||||||||||||||||||||||| |||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFYSPSPEPHLGSPSQPTSTAVLMPW 120
          |||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGIRSLCAQAQEGAGHSPSFYSPSPEEPHLGSPSQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQSSPNAPTST 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQSSPNAPTST 180

Qy    181 AVLGSWSGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERREREREKSBKGRPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||         .|||||||
Db    181 AVLGSWSGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERREREKSEKGRPWAAVVGT 240

Qy    241 PPQAHTSLQPHHIPWEPSVRESLCSTWPWKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHIPWEPSVRESLCSTWPWKNEDFNQKTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQYKEAWRGQLYGDR 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGYDEPGWVLQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGYDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVASNKELWALCLVPWVSWLACTCLMQOMKRKEKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVASNKELWALCLVPWVSWLACTCLMQOMKRKEKL 540

Qy    541 TLTSKTTTTLCLHYLAQQLQAQPLGPQLRDLCSLAAEGTWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLTSKTTTTLCLHYLAQQLQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKGHSNCIIDLEKTLEAY 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKGHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781 PTMWVLFRNVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAAVKSLCKTLRRPRCLLE 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMWVLFRNVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAAVKSLCKTLRRPRCLLE 840

Qy    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy    961 TTLSDEWRQELRALEQEKRPQLLIFSRRKPSVWMPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTLSDEWRQELRALEQEKRPQLLIFSRRKPSVWMPTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy   1021 ASHVAQANLKLLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
          ||||||||||||||||||||||||.|||||||||||||||||||||||||||||||||
Db   1021 ASHVAQANLKLLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFL 1080

Qy   1081 GPTGPVATEVVDKBRNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIFPCVWDQPFLG-EI 1139
          |||||||||||||||||||||.|||||||||||:|:|||||:::|:||:||:  .:
Db   1081 GPEGNVDVELIDKSTNRYSVWFPPTAGWILWSATGLGFLVRDEVTVTIAFGSWSQHLADL 1140

Qy   1140 NPQHSWMYAGPLLDIKARP-GAVEAVHLPHFVALQGGGHVDTSLFQMAHRKEEGMLLEKPA 1198
          ||:.:|||||||:::||| |||:||||:||||||||||||||||||:||.::||:|||:
Db   1141 QHHEQWLVGGPLFDVTAEPEEAVAEIHLPHFISLQ-GEVDGVSWFLVAIFKNEGMVLEHPA 1199

Qy   1199 RVELHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCS 1258
          |||:|||||||||||||||||:::.|||:.:||||::|||:|||:|::||||:|.||:
Db   1200 RVEPFYAVLBSPSFSLMGTLLRIASGTRLSIPITSNTLIYYHHPBDIKFHLYLVPSDAL 1259

Qy   1259 IR----------------------------------------KELELCYRSPGEDQLF 1276
          :.                                        ||.|||:|||:||:||
Db   1260 LTKAIDDEBDRFHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF 1319

Qy   1277 SEFYVGHLGSGIRLQVKDKKDEILVWEALVKPGDIMPATTLIPPARIAVPSPLDAPQLLH 1336
          |||||||||||||||:||||||::||||||||||:|||||||||||||||||||||||||
Db   1320 SKFYAGQMKEPIQLEITEKRHGTLVWDTEVKPVDLQLVAASAPP--------PFSGAA-- 1369

Qy   1337 FVDGYRBQLIARVTSVEVVLDKLH-GQVLSQBQYERVLAENTRPSQMRKLFSLSQSWDRK 1395
          |||:|:||::|||:|:||:||:|| |||||||::||||::.:.||||:||||.:|:|||:
Db   1370 FVKENHRQLQARMGDLKGVLDDLQDNEVLTENEKELVEQBKTRQSKNEALLSMVEKKGDL 1429

Qy   1396 CKDGLYQALKETHPHLIMEL 1415
          ||:::||||:||:||||:|
Db   1430 ALDVLFRSISBRDPYLVSYL 1449
```

```
RESULT 15
AAY72674
ID   AAY72674 standard; protein; 1424 AA.
XX
AC   AAY72674;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human NAC gamma or delta isoform-CARD-X1 chimeric protein.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW   chimeric protein.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FH
FT   Region          1..1048
FT                   /note= "Corresponds to 1-1048 of human NAC delta
FT                   (AAY72671) or gamma (AAY72670) isoform"
FT   Region          1049..1424
FT                   /note= "Corresponds to 56-431 of human CARD-X
FT                   protein (AAY72672)"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
```

XX
PI
XX   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   N-PSDB; AAD02765.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
     apoptosis -
XX
PS   Disclosure; Page 179-183; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC   gamma or delta isoform-CARD-X1 chimeric protein. NAC protein comprises a
CC   nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC   caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC   domain. CARD-X protein comprises a caspase-associated recruitment domain
CC   (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC   proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC   for screening modulators that modulates apoptosis, cytokine production,
CC   cytokine receptor signalling and other cellular processes. They can act
CC   as an immunogen for the production of polyclonal and monoclonal
CC   antibodies. They can also be used to diagnose and treat inflammatory
CC   disorders such as sepsis, fibrosis and arthritis and cancer pathologies
CC   such as adenocarcinomas and leukaemias.
XX
SQ   Sequence   1424 AA;

Query Match   81.9%; Score 6174; DB 22; Length 1424;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

```
Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLLALHTWEQMLGRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLLALHTWEQMLGRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKNEDDNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHHPWEPSVRESLCGSTWPWKNEDDNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTOEPRIVILOGAAGIGKSTLAROVKEAWGRGQLYGDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTOEPRIVILOGAAGIGKSTLAROVKEAWGRGQLYGDR 360

Qy    361 FQHYFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILLDGVDEPGWVLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILLDGVDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy    541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKLFSPDDLRKHGLDGAI 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGKHSNCIIDLEKTLEAY 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLYETRNKTFLTQVWAHFEEMGMCVEFDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYETRNKTFLTQVWAHFEEMGMCVEFDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781 PTMVVLFRWVPVTDAYWQILFSVLKYTRNLKEFLDLSGNSLSHSAVKSLCKTLRRPCLLE 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMVVLFRWVPVTDAYWQILFSVLKYTRNLKEFLDLSGNSLSHSAVKSLCKTLRRPCLLE 840

Qy    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLG--- 957

Qy    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTSSSLKRQRLGSERA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 ----------------KPSVWTPTEGLDTGEMSNSTSSSLKRQRLGSERA 990

Qy   1021 ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPELLLCVPSPASOQDLHTKPLGTDDDFW 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    991 ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPELLLCVPSPASOQDLHTKPLGTDDDFL 1050

Qy   1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLG-EI 1139
          :|||||||||||||||||||||||:|||:||||||||||||||||||||||||||: ::
Db   1051 GPEGNVDVELIDKSTNRYSVWFPTAGWYLWSATGLGFLVRDEVTVVTIAFGSMSQHLALDL 1110

Qy   1140 NPQHSWMVAGPFLLDIKAEP-GAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPA 1198
          :|||||:|||||:||||: ||||||||:|:||||||:|:||:||||:||||||::||||
Db   1111 QHHEQMLVGGFLPDVTAEPEEAVAEIHLPHFISLQ-GEVDVSWFLVAHFPKNEGMVLEHPA 1169

Qy   1199 RVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHYLLIPSDCS 1258
          ||:|:||||||:|:|:|||||:||:||||:|:||:||||||||||:|:||||||:||
Db   1170 RVEPFYAVLESSPFSFSLMGILLRIASGTRLSIPITSNTLIYHPHPEEDIKFHLYLVPSDAL 1229

Qy   1259 IR------------------------------------KELELCYRSPGEDOLF 1276
                                                :|||:||||||||:||
Db   1230 LTKAIDDEEDRFHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF 1289

Qy   1277 SEFYYCHLGSGIRLQVKDKDETLVWEALYKPGDLMPATTLIPPARIAVPSPIDAPQLLH 1336
          :||:|||||:||||||:||||:||||||||:||||:|||||||||||||||:||||||
Db   1290 SKFYAGQMKEPIQLEITEKRIGTLVWDTEVRKPVDLQLVAASAPP------PFSGAA-- 1339

Qy   1337 FVDQYREQLIARVTSVEVVLDKLH-GQVLSQEQYERVLAENTRPSQMRKLFSLSQSMDRK 1395
          ||||||||:||||||:|:|:|||| |:||:|||||:|||||||||:||||:|||:|||
Db   1340 FVKENHRQLQARMGDLKGVLDDLQDNEVLTENEKELVEQERTRQSKNEALLSMVEKKGDL 1399

Qy   1396 CKDGLYQALKETHPHLIMEL 1415
          :||::|||::||||||:|||
Db   1400 ALDVLFRSISISRDPYLVSYL 1419
```

QM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 03:03:41 ; Search time 1728.98 Seconds
                                        (without alignments)
                                        11475.528 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccagggcctggagaggt..........aggaataagaagttacctac 5444
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 2434939 seqs, 182278265 residues
Total number of hits satisfying chosen parameters: 4869878

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   Published_Applications NA:*
                1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                9: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                13: /cgn2_6/ptodata/1/pubpna/US0A_PUBCOMB.seq:*
                14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                17: /cgn2_6/ptodata/1/pubpna/US0_NEW_PUB.seq:*
                18: /cgn2_6/ptodata/1/pubpna/US0_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	5444	100.0	5444	10	US-09-996-617-1	Sequence 1, Appli
2	5444	100.0	5444	10	US-09-931-071-1	Sequence 1, Appli
3	5444	100.0	5444	11	US-09-956-712-3	Sequence 3, Appl
4	5284.2	97.1	6531	11	US-09-956-712-11	Sequence 11, Appl
5	4882.4	89.7	5100	11	US-09-956-712-12	Sequence 12, Appl
6	4287	78.7	4287	10	US-09-996-617-5	Sequence 5, Appli
7	4287	78.7	4287	10	US-09-931-071-5	Sequence 5, Appli
8	4148	76.2	4422	10	US-09-388-221-5	Sequence 5, Appli
9	4100	75.3	4200	10	US-09-388-221-3	Sequence 3, Appli
10	3958	72.7	4332	10	US-09-388-221-5	Sequence 5, Appli
11	3400.4	62.5	4556	10	US-09-388-221-9	Sequence 9, Appli
12	3210.4	59.0	4456	10	US-09-388-221-11	Sequence 11, Appl
13	2572.6	47.3	2657	11	US-09-895-298-22	Sequence 22, Appl
14	1710.8	31.4	96649	11	US-09-956-712-10	Sequence 10, Appl
15	694.4	12.8	720	13	US-10-027-632-150437	Sequence 150437,

ALIGNMENTS

16	694.4	12.8	720	14	US-10-027-632-150437	Sequence 150437,
17	564	10.4	564	13	US-10-029-385-2805	Sequence 2805, Ap
18	508.8	9.3	725	13	US-10-027-632-20320	Sequence 20320, A
19	508.8	9.3	725	14	US-10-027-632-20320	Sequence 20320, A
20	489.4	9.0	491	10	US-09-796-692-5036	Sequence 5036, Ap
21	489.4	9.0	491	12	US-10-057-475B-5036	Sequence 5036, Ap
22	489.4	9.0	491	12	US-10-154-884B-5036	Sequence 5036, Ap
23	489.4	9.0	491	15	US-10-040-862-5036	Sequence 5036, Ap
24	484.4	8.9	578	11	US-09-764-891-1745	Sequence 1745, Ap
25	482.8	8.9	487	15	US-10-060-036-3298	Sequence 3298, Ap
26	414.4	7.6	416	15	US-10-060-036-3299	Sequence 3299, Ap
27	399.8	7.3	3186	15	US-10-066-521-17	Sequence 17, Appl
28	398.2	7.3	3466	12	US-09-965-621-23	Sequence 23, Appl
29	398	7.3	3108	12	US-10-108-260A-718	Sequence 718, App
30	398	7.3	3108	15	US-10-407-866-23	Sequence 23, Appl
31	398	7.3	3218	12	US-09-965-621-23	Sequence 23, Appl
32	369.8	6.8	3102	13	US-10-132-967-6	Sequence 6, Appli
33	369.8	6.8	3102	14	US-10-407-866-67	Sequence 67, Appli
34	369.8	6.8	3102	14	US-10-127-516-6	Sequence 6, Appli
35	369.8	6.8	3102	14	US-10-027-629-6	Sequence 6, Appli
36	369.8	6.8	3857	13	US-10-132-967-4	Sequence 4, Appli
37	369.8	6.8	3857	14	US-10-127-516-4	Sequence 4, Appli
38	340.6	6.3	371	12	US-10-027-629-4	Sequence 4, Appli
39	318.4	5.8	322	10	US-10-242-535A-20893	Sequence 20893, A
40	318.4	5.8	322	10	US-09-796-692-2680	Sequence 2680, Ap
41	318.4	5.8	322	12	US-10-057-475B-2680	Sequence 2680, Ap
42	318.4	5.8	322	12	US-10-154-884B-2680	Sequence 2680, Ap
43	313.4	5.8	2691	15	US-10-040-862-2680	Sequence 2680, Ap
44	292.8	5.4	2524	11	US-09-965-621-58	Sequence 58, Appl
45	292.8	5.4	2524	12	US-10-407-866-58	Sequence 58, Appl

RESULT 1
US-09-996-617-1
Sequence 1, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4809)
US-09-996-617-1

Query Match   100.0%; Score 5444; DB 10; Length 5444;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 5444; Conservative   0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 GCCCCAGGGCCTGGAGAGGTTCTGAAGAAAACTGGGAGCCAGCAGCCCGGGGGGTCCACTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCCCCAGGGCCTGGAGAGGTTCTGAAGAAAACTGGGAGCCAGCAGCCCGGGGGGTCCACTCT 60

Qy   61 GGGTTCTGAAAGCCCATTCCCTGCTCTTGCGGTCTTCCTCCCACCCCACCTCTTCTCAGCCTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCCCCAGGGCCTGGAGAGGTTCTGAAGAAACCTGGGAGCCAGCAGCCCGGGGGGTCCACTCT 120

```
 61 Db  GGGTTCTTGAAAGCCCATTCCCTGCTCTGCGTCCCCACCCCACTTCTTCTTCAGCTCTT  120
                                                                            1201 Qy  TCAGAGAAAGGCAGGCCCCCATGGGCGCAGTGGTTGGTAGGAACGCGCCACACC  1260
121 Qy  GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGAGGACCCAGGAGGAGGGAAAGAAATCTCAGGAACA  180           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 Db  GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGAGTCCAGGAGGAATCTCAGGAACA  180          1201 Db  TCAGAGAAAGGCAGGCCCCCATGGGCGCAGTGGTGGTAGGAACGCGCCACACC  1260
181 Qy  CAGAACAGTGAGCGGTTGCCCACACACCCCAGGGAGAGGGGAAGAAATCTGAGGGAACA  240           1261 Qy  AGCCTTACAGCCCCACCACCACCCCATGGGGAGGCCTTCTTGTGAGGAGAGGAGACGC  1320
181 Db  CAGAACAGTGAGCGGTTGCCCACACACCCCAGCCCCATCTCCCGTCACCCTCAC  240                    ||||||||||||||||||||||||||||||||||||||||||||||||
241 Qy  CCTCCCTGCCTGGCCCTGGACCCCATCCCAGGACCCTCCCTTATCAGCTGACTTCTTCCAGT  300          1261 Db  AGCCTACAGCCCCACCACCACCCCATGGGGAGGCCTTCTTGTGAGGAGAGGAGACGC  1320
241 Db  CCTCCCTGCCTGGCCCTGGACCCCATCCCAGGACCCTCCCTTATCAGCTGACTTCTTCCAGT  300          1321 Qy  TGGCCCTGGAAAAATGAGGATTTTAACCAAAAAATTCACAGCTGCTACTTCTACAAAGA  1380
301 Qy  GTCTTGCAGGCCCCCTTCTGGGGCTTTTCCTTCCTACCCACTTCCCCTCTTAT  360                           ||||||||||||||||||||||||||||||||||||||||||||
301 Db  GTCTTGCAGGCCCCCTTCTGGGGCTTTTCCTTCCTACCCACTTCCCCTCTTAT  360                    1321 Db  TGGCCCTGGAAAAATGAGGATTTTAACCAAAAAATTCACAGCTGCTACTTCTACAAAGA  1380
361 Qy  CGGCCGTCTATCTGTAGGTTGCCCTGGGATTTATAAAACTGGGTTCGGAATGCTGAATAAGA  420          1381 Qy  CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAAGAAGCTGGGCCTGATTAGTTGAGGAG  1440
361 Db  CGGCCGTCTATCTGTAGGTTGCCCTGGGATTTATAAAACTGGGTTCGGAATGCTGAATAAGA  420                 ||||||||||||||||||||||||||||||||||||||||||||
421 Qy  GACGGTAAGAGCCAAGGCAAAGGACAGACAGCACTGTTTCTCTGCCTGATACCCTTCACCAC  480         1381 Db  CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAAGAAGCTGGGCCTGATTAGTTGAGGAG  1440
421 Db  GACGGTAAGAGCCAAGGCAAAGGACAGCACTGTTTCTCTGCCTGATACCCTTCACCAC  480             1441 Qy  AATCGAGGACATTTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCCTGGATACCCAAGAA  1500
481 Qy  CTGGGAACATCCCCCAGACACCCTTCTTAACTCCCGGGACAGAGATGGCTCGGCGGAGCCTGG  540                 ||||||||||||||||||||||||||||||||||||||||||
481 Db  CTGGGAACATCCCCCAGACACCCTTCTTAACTCCCGGGACAGAGATGGCTCGGCGGAGCCTGG  540        1441 Db  AATCGAGGACATTTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCCTGGATACCCAAGAA  1500
541 Qy  GGCCGGCCTGGTCCTGTTACTTGGAGTTCCTGGAGGAGGAGGAGGAGTTCCAGCTT  600              1501 Qy  CCTCGGCCATAGTACTGCAGGGGGCGGTGCTGGAATTGGGAAGTCAACACTGGCCAGGCAG  1560
541 Db  GGCCGGCCTGGTCCTGTTACTTGGAGTTCCTGGAGGAGGAGGAGGAGTTCCAGCTT  600                     ||||||||||||||||||||||||||||||||||||||||||
601 Qy  CTGCTTCGCCAATAAAGCGCACTCAGGAGCTCTTCGGGTGACACCCCGCTCAGCCAGAG  660           1501 Db  CCTCGGCCATAGTACTGCAGGGGGCGGTGCTGGAATTGGGAAGTCAACACTGGCCAGGCAG  1560
601 Db  CTGCTTCGCCAATAAAGCGCACTCAGGAGCTCTTCGGGTGACACCCCGCTCAGCCAGAG  660           1561 Qy  GTGAGGGAAGCCTGGGGGGAGGGCCAGCTGTTCAGGGGACCGGCTTCCAGCCAGCTC  1620
661 Qy  AAGACGAGTGGGCATGGAGGTGGCCTCGTACCTGTGGGTCAGTAATGGGGACCACCCGGGGCC  720                ||||||||||||||||||||||||||||||||||||||||||
661 Db  AAGACGAGTGGGCATGGAGGTGGCCTCGTACCTGTGGGTCAGTAATGGGGACCACCCGGGGCC  720        1561 Db  GTGAGGGAAGCCTGGGGGGAGGGCCAGCTGTTCAGGGGACCGGCTTCCAGCCAGCTC  1620
721 Qy  TGGGACCTTAGCCCTTCCATACCTGGGGAGCAGATGGGGGTCAGTGTGCGCCCAAGCC  780            1621 Qy  TTCAGCTTGCAGAGGAGCTGGGCCCCAGTCCAAGGTGGTGGCTGGAAGTCTGCTGGAAAA  1680
721 Db  TGGGACCTTAGCCCTTCCATACCTGGGGAGCAGATGGGGGTCAGTGTGCGCCCAAGCC  780                    ||||||||||||||||||||||||||||||||||||||||||
781 Qy  CAGGAAGGGGCAGGCCACTCTCCCTCCATTCCCCTACAGCCCAGTGAACCCCACCACTGGGGG  840        1621 Db  TTCAGCTTGCAGAGGAGCTGGGCCCCAGTCCAAGGTGGTGGCTGGAAGTCTGCTGGAAAA  1680
781 Db  CAGGAAGGGGCAGGCCACTCTCCCTCCATTCCCCTACAGCCCAGTGAACCCCACCACTGGGGG  840        1681 Qy  GATGGGACAGCCACTCCGGCTCCGGCTCCAGTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTC  1740
841 Qy  TCTCCCAGCCAACCCACTCCACACGCAGCCCCAGTTGCTAATGCCCTGGATCCATGAAATTGCCGCGG  900              ||||||||||||||||||||||||||||||||||||||||||
841 Db  TCTCCCAGCCAACCCACTCCACACGCAGCCCCAGTTGCTAATGCCCTGGATCCATGAAATTGCCGCGG  900        1681 Db  GATGGGACAGCCACTCCGGCTCCGGCTCCAGTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTC  1740
901 Qy  GGGTGCACCCAGGGCTCAGAGAGGAGGGTTTTGAGACAGCTGGGGTGTGCTCCGCCCAAGCC  960         1741 Qy  TTCATCCTCGATGGTGTTAGATGAGCCCAGGGTTCTTGGCAGGAGCCGAGTTCTGAGCTC  1800
901 Db  GGGTGCACCCAGGGCTCAGAGAGGAGGGTTTTGAGACAGCTGGGGTGTGCTCCGCCCAAGCC  960                  ||||||||||||||||||||||||||||||||||||||||||
961 Qy  CGCTGGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCAGACCCAT  1020       1741 Db  TTCATCCTCGATGGTGTTAGATGAGCCCAGGGTTCTTGGCAGGAGCCGAGTTCTGAGCTC  1800
961 Db  CGCTGGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCAGACCCAT  1020       1801 Qy  TGTCTGCACTGGAGCCAGCCACAGCCGGGAGGATCACTGCTGGGCAGTTTGCTCATCGGAAA  1860
1021 Qy  GAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCACATCCACAACGAGCCTGG  1080                      ||||||||||||||||||||||||||||||||||||||||||
1021 Db  GAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCACATCCACAACGAGCCTGG  1080               1801 Db  TGTCTGCACTGGAGCCAGCCACAGCCGGGAGGATCACTGCTGGGCAGTTTGCTCATCGGAAA  1860
1081 Qy  GGATCCCCACCCTCAGCCTTAGCACCCCAGCCTTGGGACCCAATGG  1140                      1861 Qy  ACTATACTTCCCGAGGCATCCTTCCTGATCACAGGGCACAGCTCTGCAGAACCTC  1920
1081 Db  GGATCCCCACCCTCAGCCTTAGCACCCCAGCCTTGGGACCCAATGG  1140                             ||||||||||||||||||||||||||||||||||||||||||
1141 Qy  CCTCTGGATGAATGAAAGGTCAGGAATTTACTACACAGAAATCAGAGAGAGAGAGAGAGAAA  1200     1861 Db  ACTATACTTCCCGAGGCATCCTTCCTGATCACAGGGCACAGCTCTGCAGAACCTC  1920
1141 Db  CCTCTGGATGAATGAAAGGTCAGGAATTTACTACACAGAAATCAGAGAGAGAGAGAGAGAAA  1200     1921 Qy  ATTCCTTCTTTGGACAGGCAGCCACGTTGGGTGAGGTTCCTGGAGTCGATGCAGG  1980
                                                                                             ||||||||||||||||||||||||||||||||||||||||||
                                                                          1921 Db  ATTCCTTCTTTGGACAGGCAGCCACGTTGGGTGAGGTTCCTGGAGTCGATGCAGG  1980
                                                                          1981 Qy  AAGGAAATATTTCTACAGATATTTCACAGATGAAAAGGCAAGCAAATTAGAGCCTTTAGGTTG  2040
                                                                                      ||||||||||||||||||||||||||||||||||||||||||
                                                                          1981 Db  AAGGAAATATTTCTACAGATATTTCACAGATGAAAAGGCAAGCAAATTAGAGCCTTTAGGTTG  2040
                                                                          2041 Qy  GTCAAATCAAACAAAAGAGCTCTGGGCCCCTGGTGTCTTGGTGGTCCTGGCTGGGCC  2100
                                                                                      ||||||||||||||||||||||||||||||||||||||||||
                                                                          2041 Db  GTCAAATCAAACAAAAGAGCTCTGGGCCCCTGGTGTCTTGGTGGTCCTGGCTGGGCC  2100
                                                                          2101 Qy  TGCACTTGCCTGATCGCAGCGAAGCGCAGAAAAACTCACAGTCGACTTCCAAGACC  2160
                                                                                      ||||||||||||||||||||||||||||||||||||||||||
                                                                          2101 Db  TGCACTTGCCTGATCGCAGCGAAGCGCAGAAAAACTCACAGTCGACTTCCAAGACC  2160
                                                                          2161 Qy  ACCAACAACCCTTCTCGTTCTACATTACCTTGCCCAGGCTCTCCAAGTCCAGGACCC  2220
                                                                                      ||||||||||||||||||||||||||||||||||||||||||
                                                                          2161 Db  ACCAACAACCCTTCTCGTTCTACATTACCTTGCCCAGGCTCTCCAAGTCCAGGACCC  2220
                                                                          2221 Qy  CAGCTCAGAGAGACCTTCTGCTTCTGGCTGCTGAGGGGCATCTGGGCAAAAAAAGAGACCCTTTTTC  2280
                                                                                      ||||||||||||||||||||||||||||||||||||||||||
                                                                          2221 Db  CAGCTCAGAGAGACCTTCTGCTTCTGGCTGCTGAGGGGCATCTGGGCAAAAAAAGAGACCCTTTTTC  2280
```

```
Qy  2281 AGTCCAGATGACCTCAGGAGACCATGGGTTAGATGGGGGGCCAGCATCACATCTCCACTTTCTTGAAG 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 AGTCCAGATGACCTCAGGAGAACCATGGGTTAGATGGGGGCCATCATCTTCCACTTTCTTGAAG 2340

Qy  2341 ATGGGTATTCTTCAAGAGACACCCCATCCCTCCTGAGCTACAGCTTCACTTCTTTTC 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 ATGGGTATTCTTCAAGAGACACCCCATCCCTCCTGAGCTACAGCTTCACTTCTTTTC 2400

Qy  2401 CAAGAGTTCTTTGCAGCAATGTCCTATGTCTTTGGAGGATGAGAGGGGAGGTAAACAT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CAAGAGTTCTTTGCAGCAAATGTCCTATGTCTTTGGGAGGATGAGAGGGGAGGTAAATAGC 2460

Qy  2461 TCTAATTGCATCATAGATTTGGAAAAAGAACGCTAGAAGCATTAGATTGGAAAAACATGGCCCTGTTT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GGGGCATCAACCACACAGTTTCTTATTGGGCCTAGTTAAGTCAGGGGGAGAGAGAGATG 2520

Qy  2521 GGGGCATCAACCACACAGTTTCTTATTGGGCCTGTTAAGTCAGGGGGAGAGAGAGATG 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 GAGAACATCTTTCACTGCCGGGCTGTCTCAGGGGGAGGAAACCTTGATGCAGTGGGTTCCCGTCC 2580

Qy  2581 GAGAACATCTTTCACTGCCGGGCTGTCTCAGGGGGAGGAAACCTTGATGCAGTGGGTTCCCGTCC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 CTGCAGCTGCTGTGCTGAGCCACACTTCTCTGAGTCCCTCCACTGGTAGTACGAGACTCGG 2640

Qy  2641 CTGCAGCTGCTGTGCTGAGCCACACTTCTCTGAGTCCCTCCACTGGTAGTACGAGACTCGG 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 AACAAAAACGTTCCTGACACAAGTGATGGCCCATTTCGGAAGAAATGGGCATGTGTGTTAGAA 2700

Qy  2701 AACAAAAACGTTCCTGACACAAGTGATGGCCCATTTCGGAAGAAATGGGCATGTGTGTTAGAA 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 ACAGACACATGGAGCTCTTAGTGTGCACTTTCTGCATTAAATTCAGCCCGCCCACGTGAAGAAG 2760

Qy  2761 ACAGACACATGGAGCTCTTAGTGTGCACTTTCTGCATTAAATTCAGCCCGCCCACGTGAAGAAG 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 CTTCAGCTGATTGAGGGCCAGCGAGCAGCATCAACATCGGACCCCACCCCAGGTGAGTACTGC 2820

Qy  2821 CTTCAGCTGATTGAGGGCCAGCGAGCAGCATCAACATCGGACCCCACCCCAGGTAGTACTGC 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCTCCGGTCTCAAGGTC 2880

Qy  2881 TTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCTCCGGTCTCAAGGTC 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 ACCAGAAAACCTGAAGGAGCTCGGAACCTTAAGTGGAAAACTCGCTGAGCCTGAGCTGAAG 2940

Qy  2941 ACCAGAAAACCTGAAGGAGCTCGGAACCTTAAGTGGAAAACTCGCTGAGCCTGAGCTGAAG 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 AGTCTTTTGTTAAGACCCTGAGGACGCCCCTCGCTGCCTCCTCCTGGTTGGCTGGTGGC 3000

Qy  3001 AGTCTTTTGTTAAGACCCTGAGGACGCCCCTCGCTGCCTCCTCCTGGTTGGCTGGTGGC 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TGTGGGCCTCACAGCTGAGGAGACTCAAGGACCTTGCCTTTGGGCTGAGAGCCAACCAGACC 3060

Qy  3061 TGTGGGCCTCACAGCTGAGGAGACTCAAGGACCTTGCCTTTGGGCTGAGAGCCAACCAGACC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 CTGACCGAGCTGGACCTGAGCTTCAAGTGTCTCACCGGATGCTGGGACCCAAACACCTTTGC 3120

Qy  3121 CTGACCGAGCTGGACCTGAGCTTCAAGTGTCTCACCGGATGCTGGGACCCAAACACCTTTGC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CAGAGACTGAGACACGCGGAGCTTGCAAGCTACAGCGGACTCTGCAGCTGTGTGCCCTC 3180

Qy  3181 CAGAGACTGAGACACGCGGAGCTTGCAAGCTACAGCGGACTCTGCAGCTGTGTGCCCTC 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 ACGTCTTGACTCGCTGCCAGGACCTGGCCTCTGTGTTGCTTAGTGCCAGGCTGAAAGGAG 3240

Qy  3241 ACGTCTTGACTCGCTGCCAGGACCTGGCCTCTGTGTTGCTTAGTGCCAGGCTGAAAGGAG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 CTAGACCTGCAGCAGAGAACAAACCTGGATGACGTTGGGGTGCGACTGCTCTGTGAGGGGGTC 3300

Qy  3301 CTAGACCTGCAGCAGAGAACAAACCTGGATGACGTTGGGGTGCGACTGCTCTGTGAGGGGGTC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CTAGACCTGCAGCAGAGGAACAAACCTGGATGACGTTGGGGCCTGCGGACTGCTCTGTGAGGGGCTC 3360

Qy  3361 AGGCATCCTGCCTGCAAACTCATACCGCCTGGGGGCTGGACCAGAGACAACTCTGAGTGAATGAG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 AGGCATCCTGCCTGCAAACTCATACCGCCTGGGGGCTGGACCAGAGACAACTCTGAGTGAATGAG 3420

Db  3421 ATGAGGGCAGGAACTGAGGGCCCTGGACGCAGGAGGAGAAACCTCAGCTGCTGCTCATCTTCAGCAGA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3421 ATGAGGGCAGGAACTGAGGGCCCTGGACGCAGGAGGAGAAACCTCAGCTGCTGCTCATCTTCAGCAGA 3480

Qy  3481 CGGAAACCAAGTGTGATGACCCCCTACTGAGGGCCTGGATACGGGAGGAGGAGAGGAGGAA 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CGGAAACCAAGTGTGATGACCCCCTACTGAGGGCCTGGATACGGGGAGATGAGCTAATAGC 3540

Qy  3541 ACATCCTCACTCAAGCGGGCCAGAGAACTCGGATCAGAGACGGGGCGGCCTTCCCAGTTGCTCAG 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 ACATCCTCACTCAAGCGGGCCAGAGAACTCGGATCAGAGACGGGGCGGCCTTCCCATGTTGCTCAG 3600

Qy  3601 GCTAAATCTCAAAACTCCTGGGACGTGAGCAGAGAGATCTTCCCAAATTGCTGAGATTGCTCAAGGG 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 GCTAAATCTCAAAACTCCTGGGACGTGAGCAGAGAGATCTTCCCAAATTGCTGAGATTGCTCAAGG 3660

Qy  3661 AGCTCCCCAGAGGTAGTACGGGTGGGAACTCTTGTGGTGGGTGCCTTCTCCTGCTCTTCAAGGG 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 AGCTCCCCAGAGGTAGTACGGGTGGGAACTCTTGTGGTGGGTGCCTTCTCCTGCTCTTCAAGGG 3720

Qy  3721 GACCTGCATACAGAAAGCCTTTGGGGGACTGACGATGACTTCTGTGCCTGGATG 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GACCTGCATACAGAAAGCCTTTGGGGGACTGACGATGACTTCTGTGCCTGGATG 3780

Qy  3781 GCTACTGAGGTAGTTGACAAAAGAAAAGAAACTTGTACGGAGTTCACTTCGCTGGC 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GCTACTGAGGTAGTTGACAAAAGAAAGAAACTTGTACGGAGTTCACTTCTGCGGC 3840

Qy  3841 TCCTACCGCTGGCCAACACGGGGTCTCTGCTTTGTGGATGAGGAAGCGGTTGACCGGTTGAG 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TCCTACCGCTGGCCAACACGGGTCTCTGCTTTGTGGATGAGGAAGCGGTGACCGGTTGACCCCAC 3900

Qy  3901 ATTGAATTCTGTGTGTGGGGACCAGTTCCTGGGGTGAGGATCAACCACAGCACGCTGGATG 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 ATTGAATTCTGTGTGTGGGGGACCAGTTCCTGGGGTGAGGATCAACCACAGCACGCTGGATG 3960

Qy  3961 GTGGCAGGGGCTCTGCTGGACATCAAGGCTGAGCCTGAGGTCGTGTGCACCTC 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GTGGCAGGGGCTCTGCTGGACATCAAGGCTGAGCCTGAGGTCGTGTGCACCTC 4020

Qy  4021 CCTCACTTTGTGGCCTCTCCAAGGGGGCCATGTGGGACACATCCCTGTTCCAAATGGCCCAC 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CCTCACTTTGTGGCCTCTCCAAGGGGGCCATGTGGGACACATCCCTGTTCCAAATGGCCCAC 4080

Qy  4081 TTTAAAGAGGAGGGGAGAGCTCCTGGGAGAAGCCAGCCAGGGTGGAGCGCATCACATAGTT 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 TTTAAAGAGGAGGGGAGAGCTCCTGGGAGAAGCCAGCCAGGGTGGAGCGCATCACATAGTT 4140

Qy  4141 CTGGAAAACCCCAGAGCTTCTCCCCCTCGGGGAGTCCTCCTTGAAAATGATCCATAATGCCCTG 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 CTGGAAAACCCCAGAGCTTCTCCCCCTCGGGGAGTCCTCCTTGAAAATGATCCATAATGCCCTG 4200

Qy  4201 CGCTTCATTCCCGTCACCTCTGTGTGTTGCTTTACCACGCGGGTCCATCCTGAGGAAGTC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 CGCTTCATTCCCGTCACCTCTGTGTGTTGCTTTACCACGCGGGTCCATCCTGAGGAAGTC 4260

Qy  4261 ACCTTCCACCTCTACCTGATTCCCAAGTGACTGCTCCAAGTGACTGGAGCTCTGC 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 ACCTTCCACCTCTACCTGATTCCCAAGTGACTGCTCCAAGTGACTGGAGCTCTGC 4320

Qy  4321 TATCGAAGCCCTGGAGAAGACCAGCTGTTCTCGGGAGTTCTACGTTGGCCACTTGGGGATCA 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TATCGAAGCCCTGGAGAAGACCAGCTGTTCTCGGGAGTTCTACGTTGGCCACTTGGGGATCA 4380

Qy  4381 GGGGATCAGGCTGCAAGTGAAAGACAAGAGAAAGAGAGACTCTGGTGTGTGGGGAGGCCTTGGTG 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 GGGGATCAGGCTGCAAGTGAAAGACAAGAGAAAGATGAGACTCTGGTGTGTGGGGAGGCCTTGGTG 4440

Qy  4441 AAACCAGGAGGATCTCATGCCTGCAACTACTTCTGATCCCTCCAGCGCCCCCATAGCCGGTACCT 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Query Match   100.0%; Score 5444; DB 10; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[Page contains two columns of paired Db/Qy nucleotide sequence alignment lines with position numbers. Left column position markers range from 4441–5444; right column position markers range from 1–780.]

```
Db   721  TGGGACCTTAGCCCTTCCATTACCTGGGAGCAGAATGGGGGCTGAGGTGCGGCCCAAGCC
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy

Qy   781  CAGGAAGGGGGCAGGCAGCCACTCTTCCTTCCTCCATTCCCCTTACACGCCCACCTGGGG  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CAGGAAGGGGGCAGGCAGCCACTCTTCCTTCCTCCATTCCCCTTACACGCCCACCTGGGG  840

...
```

(Continued DNA sequence alignment between Qy and Db sequences, positions 721–2940)

```
Qy  2941 ACCAGAAAACCTGAAAGGAGCTGAGCCTAAGTGGAAACTGGCTGAGCCAGTCGCAGTGAAG 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 ACCAGAAAACCTGAAAGGAGCTGAGCCTAAGTGGAAACTGGCTGAGCCAGTCGCAGTGAAG 3000

Qy  3001 AGTCTTTGTAAGACCCTGAGACCGCCTCGCTCGGTTGGCTGGC 3060
Db  3001 AGTCTTTGTAAGACCCTGAGACCGCCTCGCTCGGTTGGCTGGC 3060

Qy  3061 TGTGGCCTCACAGCTGAGGACTCGCAAGGA 3120
Db  3061 TGTGGCCTCACAGCTGAGGACTCGCAAGGA 3120

Qy  3121 CTGACCCGAGCCTGGAACCTGAGCTTCACGGA 3180
Db  3121 CTGACCCGAGCCTGGAACCTGAGCTTCACGGA 3180

Qy  3181 CAGAGACTGGAGCAGACAGCCGAGCTGCAAGA 3240
Db  3181 CAGAGACTGGAGCAGACAGCCGAGCTGCAAGA 3240

Qy  3241 ACGTCTGACTGTGTGCCAGGACCTGG 3300
Db  3241 ACGTCTGACTGTGTGCCAGGACCTGG 3300

Qy  3301 CTGAGACCTGCAGCAGAGAACAA 3360
Db  3301 CTGAGACCTGCAGCAGAGAACAA 3360

Qy  3361 AGGCATCCTGCCTGCAAACTCATACG 3420
Db  3361 AGGCATCCTGCCTGCAAACTCATACG 3420

Qy  3421 ATGAGGGCAGGACTGAGGGCCC 3480
Db  3421 ATGAGGGCAGGACTGAGGGCCC 3480

Qy  3481 CGGAAAACCAAGTGTGATGAC 3540
Db  3481 CGGAAAACCAAGTGTGATGAC 3540

Qy  3541 ACATCCTCACTCAAGCGGC 3600
Db  3541 ACATCCTCACTCAAGCGGC 3600

Qy  3601 GCTAATTCTCAAACTCCTGG 3660
Db  3601 GCTAATTCTCAAACTCCTGG 3660

Qy  3661 AGCTCCCCAGAGAGGTAGTA 3720
Db  3661 AGCTCCCCAGAGAGGTAGTA 3720

Qy  3721 GACCTGCATACGGAGAGCC 3780
Db  3721 GACCTGCATACGGAGAGCC 3780

Qy  3781 GCTACTGAGGTAGTTGACA 3840
Db  3781 GCTACTGAGGTAGTTGACA 3840

Qy  3841 TCCTTACCCGCTGGCCCA 3900
Db  3841 TCCTTACCCGCTGGCCCA 3900

Qy  3901 ATTGAATTCTGTGTGTGGG 3960
Db  3901 ATTGAATTCTGTGTGTGGG 3960

Qy  3961 GTGGCAGGGCCCTCTGCT 4020
Db  3961 GTGGCAGGGCCCTCTGCT 4020

Qy  4021 CCTCACTTTGTGGCTCTCCAAGGGGGCCATGTGACACATCCCTGTTCAAAATGGCCAC 4080
Db  4021 CCTCACTTTGTGGCTCTCCAAGGGGGCCATGTGACACATCCCTGTTCAAAATGGCCAC 4080

Db  4021 CCTCACTTTGTGGCTCTCCAAGGGGGCCATGTGACACATCCCTGTTCAAAATGGCCAC 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4081 TTTAAAGAGGAGGGGGATGGTCCTGGAGAAGCCAGCCAGGGGGAGCTGCATCACATAGTT 4140
Db  4081 TTTAAAGAGGAGGGGGATGGTCCTGGAGAAGCCAGCCAGGGGGAGCTGCATCACATAGTT 4140

Qy  4141 CTGGAAAAACCCAGCTTTCTCCCCCTTGGGGAGTCCTCCTGAAAAATGATCCATAATGCCCTG 4200
Db  4141 CTGGAAAAACCCAGCTTTCTCCCCCTTGGGGAGTCCTCCTGAAAAATGATCCATAATGCCCTG 4200

Qy  4201 CGCTTCATTCCCGTCACCTCGTGTGGTGTTGCTTTACCACCGGCGTCCATCCTGAGGAAGTC 4260
Db  4201 CGCTTCATTCCCGTCACCTCGTGTGGTGTTGCTTTACCACCGGCGTCCATCCTGAGGAAGTC 4260

Qy  4261 ACCTTCCACCTCTACCTGATCCCAAGTGACTGCTCCAAGTGACTGGAACTCTGC 4320
Db  4261 ACCTTCCACCTCTACCTGATCCCAAGTGACTGCTCCAAGTGACTGGAACTCTGC 4320

Qy  4321 TATCGAAGCCCTGGAGAAGAGCAGCGTTTCTCGGGGTTCGGGGGATCA 4380
Db  4321 TATCGAAGCCCTGGAGAAGAGCAGCGTTTCTCGGGGTTCGGGGGATCA 4380

Qy  4381 GGGATCAGGCTGCAAGTGGAGAAGACAAGAAAGATGATAGGCCTTGGTG 4440
Db  4381 GGGATCAGGCTGCAAGTGGAGAAGACAAGAAAGATGATAGGCCTTGGTG 4440

Qy  4441 AAACCAGGAGAGATCTCATGGCTGGAGTCTCTGGATATCGAGGCTACCT 4500
Db  4441 AAACCAGGAGAGATCTCATGGCTGGAGTCTCTGGATATCGAGGCTACCT 4500

Qy  4501 TCACCTCTGGATGGCCCGGCAGTTCTGAGCAGTGATA 4560
Db  4501 TCACCTCTGGATGGCCCGGCAGTTCTGAGCAGTGATA 4560

Qy  4561 GCCCGAGTGACATCGGGAGGTTGTCTTGGACCAGTGCTCAGCCAG 4620
Db  4561 GCCCGAGTGACATCGGGAGGTTGTCTTGGACCAGTGCTCAGCCAG 4620

Qy  4621 GAGCAGTACGAGAGGGTGCTGGGCTGAGAACACAGAGGCCACCAGTGTTC 4680
Db  4621 GAGCAGTACGAGAGGGTGCTGGGCTGAGAACACAGAGGCCACCAGTGTTC 4680

Qy  4681 AGCTTGAGCCAGTCCTGGACCGGAAGTGCAAAACAAGTCTACCAAGGAG 4740
Db  4681 AGCTTGAGCCAGTCCTGGACCGGAAGTGCAAAACAAGTCTACCAAGGAG 4740

Qy  4741 ACCCATCCTCACCTCATTATGGAACTCTGCCA 4800
Db  4741 ACCCATCCTCACCTCATTATGGAACTCTGCCA 4800

Qy  4801 CTCAGCAGCTGAGTATCAACACCAGCCCTTGACCC 4860
Db  4801 CTCAGCAGCTGAGTATCAACACCAGCCCTTGACCC 4860

Qy  4861 TTCTTTGGGTCTCAGTTCCATCTGGGTTCCAGCA 4920
Db  4861 TTCTTTGGGTCTCAGTTCCATCTGGGTTCCAGCA 4920

Qy  4921 CTAAAGTAATGGAACTTTGATGAGCCTTTGCTTCATGCCAGGGATG 4980
Db  4921 CTAAAGTAATGGAACTTTGATGAGCCTTTGCTTCATGCCAGGGATG 4980

Qy  4981 CCACAGGGGGCCCCAGTCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
Db  4981 CCACAGGGGGCCCCAGTCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040

Qy  5041 GGCAAGAGACCCCTGCAGACCTCATAGAGCCTCATGGTGCCCACAGCAAGCCTAGA 5100
Db  5041 GGCAAGAGACCCCTGCAGACCTCATAGAGCCTCATGGTGCCCACAGCAAGCCTAGA 5100

Qy  5101 GCCCTCCGGATCCCATCCAGGCGGCAAAAGAGGAATAGGAGGGACATGGAACCATTTGCCTC 5160
Db  5101 GCCCTCCGGATCCCATCCAGGCGGCAAAAGAGGAATAGGAGGGACATGGAACCATTTGCCTC 5160
```

RESULT 3

US-09-956-712-3

Sequence 3, Application US/09956712
Publication No. US20030092648A1

GENERAL INFORMATION:
  APPLICANT: C. Frank Bennett
  APPLICANT: Susan M. Freier
  TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
  FILE REFERENCE: RTS-0326
  CURRENT APPLICATION NUMBER: US/09/956,712
  CURRENT FILING DATE: 2001-09-19
  NUMBER OF SEQ ID NOS: 91

  SEQ ID NO 3
  LENGTH: 5444
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (523)...(4812)

Query Match   100.0%;   Score 5444;   DB 11;   Length 5444;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 5444; Conservative 0;

[DNA sequence alignment between Query (Qy) and Database (Db), positions 1–5444, with 100% identity match lines shown between each aligned pair.]

```
Qy 3661 AGCTCCCCAGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCGTGCCTCTTCTCCTCCAAGGG 3720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 AGCTCCCCAGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCGTGCCTCTTCTCCTCCAAGGG 3720

Qy 3721 GACCTGCATACGAAGACCTTTTGGGGACTGACGATGACTCTCTGGGGGCCCCTGTG 3780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 GACCTGCATACGAAGACCTTTTGGGGACTGACGATGACTCTCTGGGGGCCCCTGTG 3780

Qy 3781 GCTACTGAGGTAGTTGACAAAGAAAAGAAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GCTACTGAGGTAGTTGACAAAGAAAAGAAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840

Qy 3841 TCCTACCCGGCTGGGCCCAACACGGGGTTCTCGCTTTTGTGATGAAGAGAGCGCGTTGAG 3900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TCCTACCCGGCTGGGCCCAACACGGGGTTCTCGCTTTTGTGATGAAGAGAGCGCGTTGAG 3900

Qy 3901 ATTGAATTCTGTGTGTTGGGGAGGAGGACCAGTTCCTGGGGTGAGATCAAACCCACAGACAGCTGGATG 3960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 ATTGAATTCTGTGTGTTGGGGAGGAGGACCAGTTCCTGGGGTGAGATCAAACCCACAGACAGCTGGATG 3960

Qy 3961 GTGGCAGGGCCCTTGCTGGACATCAAAGGCTGAGCCTGGAAGCTGTGTGCACCTC 4020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GTGGCAGGGCCCTTGCTGGACATCAAAGGCTGAGCCTGGAAGCTGTGTGCACCTC 4020

Qy 4021 CCTCACTTTGTTGGCTCTCCCAAGGGGGGCCCATGTGGACTGAGCCTGGACTGTTGAG 4080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CCTCACTTTGTTGGCTCTCCCAAGGGGGGCCCATGTGGACTGAGCCTGGACTGTTGAG 4080

Qy 4081 TTTTAAAGAGGGAGGGGATGGCTTCCTGGAGAAGCCAGCCAGGGGTGGAGCTGCATCACATAGTT 4140
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 TTTTAAAGAGGGAGGGGATGGCTTCCTGGAGAAGCCAGCCAGGGGTGGAGCTGCATCACATAGTT 4140

Qy 4141 CTGGAAAAACCCCAGCTTCTCCCCTTGGGAGTCCTCCTGAAAATGATCATATAATGCCCTG 4200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 CTGGAAAAACCCCAGCTTCTCCCCTTGGGAGTCCTCCTGAAAATGATCATATAATGCCCTG 4200

Qy 4201 CGCTTCATTCCCGTCACCTCTGTGGTGTTGCTTTTACCACCGCGGTTCCATCCTGAGGAAGTC 4260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4201 CGCTTCATTCCCGTCACCTCTGTGGTGTTGCTTTTACCACCGCGGTTCCATCCTGAGGAAGTC 4260

Qy 4261 ACCTTCCACCTCTTACCTGATCCCAAGTGACTGCTCCATTCGGAAGGAACTGGAGCTCTGC 4320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 ACCTTCCACCTCTTACCTGATCCCAAGTGACTGCTCCATTCGGAAGGAACTGGAGCTCTGC 4320

Qy 4321 TATCGAAGCCCTGGAGAAGACCAGCTGTTCTACGTTGGCCACTTGGGATCA 4380
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TATCGAAGCCCTGGAGAAGACCAGCTGTTCTACGTTGGCCACTTGGGATCA 4380

Qy 4381 GGGATCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTCGTGGTGTGGAGGGGCCTTGGTG 4440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 GGGATCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTCGTGGTGTGGAGGGGCCTTGGTG 4440

Qy 4441 AAACCAGGAGATCTCATGCCTGCAACTACTCTGATCCCTCCAGCCCGCCATAGCCGTACCT 4500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4441 AAACCAGGAGATCTCATGCCTGCAACTACTCTGATCCCTCCAGCCCGCCATAGCCGTACCT 4500

Qy 4501 TCACCTCTGGATGCCCCGAGTTGTTGGTGGACCAGTATCGAGAGCAGCTGATA 4560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TCACCTCTGGATGCCCCGAGTTGTTGGTGGACCAGTATCGAGAGCAGCTGATA 4560

Qy 4561 GCCCGAGTGACATGGTGGTGGAGGTTGTCTTGGACAAAACTGCATGGACAGTGTGCTGAGCCAG 4620
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 GCCCGAGTGACATGGTGGTGGAGGTTGTCTTGGACAAAACTGCATGGACAGTGTGCTGAGCCAG 4620

Qy 4621 GAGCAGTACGAGAGGGTGCTGGGCTGAGAACACGAGGCCCAGATGCGGGAAGCTGTTC 4680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4621 GAGCAGTACGAGAGGGTGCTGGGCTGAGAACACGAGGCCCAGATGCGGGAAGCTGTTC 4680

Qy 4681 AGCTTGAGCCAGTGTCCTGGGACCGGAAGTGCAAAGATGGACTCTTACCCAAGCCCTGAAAGGAG 4740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4681 AGCTTGAGCCAGTGTCCTGGGACCGGAAGTGCAAAGATGGACTCTTACCCAAGCCCTGAAAGGAG 4740

Qy 4741 ACCCATCCTTCACCTCATTATGTGGAACTCTGGGAGAAGGGCAGCAAAAAGGGGACTCCTGCCA 4800
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4741 ACCCATCCTTCACCTCATTATGTGGAACTCTGGGAGAAGGGCAGCAAAAAGGGGACTCCTGCCA 4800

Qy 4801 CTCAGCAGCTGAAGTATCAACACACAGCCAGCCCTTTGACCCTTGAGTCGTGGCTTGTTGGGCTGACCC 4860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4801 CTCAGCAGCTGAAGTATCAACACACAGCCAGCCCTTTGACCCTTGAGTCGTGGCTTGTTGGGCTGACCC 4860

Qy 4861 TTCTTTGGGTCTCAGTTTCTTTCTCTGCAAACAAGTTGCCATCTGGGTTTGCCTTCCAGCA 4920
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4861 TTCTTTGGGTCTCAGTTTCTTTCTCTGCAAACAAGTTGCCATCTGGGTTTGCCTTCCAGCA 4920

Qy 4921 CTAAAGTAATGGAACTTTGATGATGCCCTTTGCTGTGTCTCCAGTGCCCAGGGATG 4980
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4921 CTAAAGTAATGGAACTTTGATGATGCCCTTTGCTGTGTCTCCAGTGCCCAGGGATG 4980

Qy 4981 CCACAGGGGGGCCCCAGTCCAGGTGGCCTAACAGCAATCTCAGGGAATGTCCATCTCGGAGCT 5040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4981 CCACAGGGGGGCCCCAGTCCAGGTGGCCTAACAGCAATCTCAGGGAATGTCCATCTCGGAGCT 5040

Qy 5041 GGCAAGACCCCTGCAGAACCTCATCAGACGCCTCATCTCGGTGGCCACAGCAGCCAAGCCTAGA 5100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5041 GGCAAGACCCCTGCAGAACCTCATCAGACGCCTCATCTCGGTGGCCACAGCAGCCAAGCCTAGA 5100

Qy 5101 GCCCTCCGGAGATCCCATCCAGGCGCGAAAGAGGAAAATTAGGAGGGGACATGGAACCCATTTGCCTC 5160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5101 GCCCTCCGGAGATCCCATCCAGGCGCGAAAGAGGAAAATTAGGAGGGGACATGGAACCCATTTGCCTC 5160

Qy 5161 TGGCCTGTGTCACAGGGTGAGCCCCAAAAATTGGGGGTTCAGCGTGGGGGGAGGCCACGTGGATTC 5220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5161 TGGCCTGTGTCACAGGGTGAGCCCCAAAAATTGGGGGTTCAGCGTGGGGGGAGGCCACGTGGATTC 5220

Qy 5221 TTGGGCTTTGTACAGGAGATCTACAAGAGGCAAGCCAACAGAGTAAAGTGGAAGGAAGTTT 5280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5221 TTGGGCTTTGTACAGGAGATCTACAAGAGGCAAGCCAACAGAGTAAAGTGGAAGGAAGTTT 5280

Qy 5281 ATTCAGAAAATAAAGGAGTATCACAGCTCTCTTTTAGAATTTGTCTAGCGACGCTTTCCAGTT 5340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5281 ATTCAGAAAATAAAGGAGTATCACAGCTCTCTTTTAGAATTTGTCTAGCGACGCTTTCCAGTT 5340

Qy 5341 TTTACCAGAAAAACCCCTATAAAATTAAAAATTTTTTACTTAAATTTAAAATTAAAAAAAT 5400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5341 TTTACCAGAAAAACCCCTATAAAATTAAAAATTTTTTACTTAAATTTAAAATTAAAAAAAT 5400

Qy 5401 ACAAAAAAAAGAAAAAAATGAAAAATAAAAGGAATAAGAAAGAAGTTACCTAC 5444
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5401 ACAAAAAAAAGAAAAAAATGAAAAATAAAAGGAATAAGAAAGAAGTTACCTAC 5444
```

Query Match   97.1%;   Score 5284.2;   DB 11;   Length 6531;
Best Local Similarity   97.5%;   Pred. No. 0;
Matches 5441;   Conservative 0;   Mismatches 3;   Indels 135;   Gaps 2;

```
Qy 1 GCCCCAGGGCCTGAGAGGTTCTGAAGAAACCTGGGAGCGCAGCCCGCGGGGCTCCACTCT 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1  GCCCCAGGGCCTGAGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCCCGGCGGGGCTCCACTCT  60

Qy     61  GGGTTCTGAAAGCCCATTCCCTCGCTCCTCTGCTCTCTTCTTCTCAGCCTT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GGGTTCTGAAAGCCCATTCCCTGCTCTCGCGGGCTCCTCCACCCCACCTCTTCTTCTCAGCCTT  120

Qy    121  GCAGCTCAAGGGTTGATCTTCAGGAGTCCAGGAGCCAGGAGAGGGGAAGGAGAATCTGAGGAACA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GCAGCTCAAGGGTTGATCTTCAGGAGTCCAGGAGCCAGGAGAGGGGAAGGAGAATCTGAGGAACA  180

Qy    181  CAGAACAGTGAGCGGTTGCCCACCACACCCCATCTCCCCTCACCCTCAC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CAGAACAGTGAGCGGTTGCCCACCACACCCCATCTCCCCTCACCCTCAC  240
```

Nucleotide sequence alignment (Query "Qy" vs. Database "Db"), residues 2221–4305.

RESULT 5
US-09-956-712-12
; Sequence 12, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
; FILE REFERENCE: RTS-0326
; CURRENT APPLICATION NUMBER: US/09/956,712
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 12
; LENGTH: 5100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)...(4449)
US-09-956-712-12

Query Match   89.7%;   Score 4882.4;   DB 11;   Length 5100;
Best Local Similarity   99.5%;   Pred. No. 0;   Mismatches 26;   Indels 0;   Gaps 0;
Matches 4898;   Conservative 0;   Mismatches 26;   Indels 0;   Gaps 0;

```
Db   753 GCAGGAGGCTCCTCGGACCAAATGCCTCTCGGATGAAACGTCAGGAATTTACTACACAGA 812

Qy  1176 AATCAGAGAAAGAGAGAGAGAAAATCAGAGAAAGGCAGGCCCCATGGGCAGGCGGTGGT 1235
Db   813 AATCAGAGAAAGAGAGAGAGAAAATCAGAGAAAGGCAGGCCCCATGGGCAGGCGGTGGT 872

Qy  1236 AGGAACGGCCCCCACAGGCGGCCACACAGCCTACAGCCCCACCACCACCCCATGGGGACCCTTC 1295
Db   873 AGGAACGGCCCCCACAGGCGGCCACACAGCCTACAGCCCCACCACCACCCCATGGGGAGCCTTC 932

Qy  1296 TGTGAGAGAGGAGCCTTCGTTCCACATGGCCCTGGAAAAATGAGGATTTTAACCAAAAATTT 1355
Db   933 TGTGAGAGGAGCCCCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAACCAAAAATT 992

Qy  1356 CACACAGCTGCTACTTTCTACAAAAGACCTCACCCCAGGAAGCCAAGATCCCCTGGTCAAGAG 1415
Db   993 CACACAGCTGCTACTTTCTACAAAAGACCTCACCCCAGGAAGCCAAGATCCCCTGGTCAAGAG 1052

Qy  1416 AAGCTGGCCTGATTTATGTGAGGAGGAACATTTAATTGAGATCAGAGACTTATT 1475
Db  1053 AAGCTGGCCTGATTTATGTGAGGAGGAACATTTAATTGAGATCAGAGACTTATT 1112

Qy  1476 TGGGCCCAGGGCCTGGATACCCAAGAACCTCGGCATAGTCATACTGCAGGGGGGCTGGAAT 1535
Db  1113 TGGGCCCAGGCCTGGATACCCAAGAACCTCGGCATAGTCATACTGCAGGGGGGCTGGAAT 1172

Qy  1536 TGGGAAGTCAACACTGGGCAGGCAGGTGAAGGGACCCTGGGGGAGAGGGGGGCTGTATGG 1595
Db  1173 TGGGAAGTCAACACTGGGCAGGCAGGTGAAGGGACCCTGGGGGAGAGGGGGGCTGTATGG 1232

Qy  1596 GGACCGGCTTCCAGCATGTCTTCTACTTCAGCTGCAGAGAGTCCCAGTGTCCAAGGTGGT 1655
Db  1233 GGACCGGCTTCCAGCATGTCTTCTACTTCAGCTGCAGAGAGTCCCAGTGTCCAAGGTGGT 1292

Qy  1656 GAGTTCTGCTGAGTCTCATCGGAAAAAGATGGGACCAGCCACTCCGGCTCCATTAGACAGAT 1715
Db  1293 GAGTTCTGCTGAGTCTCATCGGAAAAAGATGGGACCAGCCACTCCGGCTCCATTAGACAGAT 1352

Qy  1716 CCTGTCTTAGCCCAGAGCGCGCTGCTTGCTTCTTCATCCTCGATGGTAGAGATGAGCCAGAGATGGGGT 1775
Db  1353 CCTGTCTTAGCCCAGAGCGCGCTGCTTGCTTCTTCATCCTCGATGGTAGAGATGAGCCAGAGATGGGT 1412

Qy  1776 CTTGCAGGAGCCGGAGTTCTGTGTCTGCACTGGGAGCTCAGCCAGCCCACGCGGCGGATGC 1835
Db  1413 CTTGCAGGAGCCGGAGTTCTGTGTCTGCACTGGGAGCTCAGCCAGCCCACGCGGCGGATGC 1472

Qy  1836 ACTGGCTGGGGCAGTTTGCTGGGGGAAAACTATACTTCCTGATCACCGGC 1895
Db  1473 ACTGGCTCGGCGGAGTTTGCTGGGGGAAAAACTATACTTCCCGATCACCGGC 1532

Qy  1896 TCGGACCCAGCCACAGCCTCTGCAGAAACCTCCATTCCTTCTTTCTTTTGGGACAGGGCAGGGGTTAGAGGT 1955
Db  1533 TCGGACACCAGCTCTGCAGAACCTCATTCCTTCTTTCTTTTGGGTAGAGGT 1592

Qy  1956 CCTGGGGGTTCTCTCAGCAGCAAGGAATATATTTTCTACAGATTGAAAG 2015
Db  1593 CCTGGGGTTCTCTGAGTCCAGCAGGAATATTTTTTCACAGATTGAAAG 1652

Qy  2016 GCAAGCAATTAGAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGAGCTTCTGGGCCCCTGTGTCT 2075
Db  1653 GCAAGCAATTAGAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGAGCTTCTGGGCCCCTGTGTCT 1712

Qy  2076 TGTGCCCTGGGTGTCCTGGGCTGGCCTGCACTTGCCTGATCGCAGACGATGAAAGCGAAAGGA 2135
Db  1713 TGTGCCCTGGGTGTCCTGGGCTGGCCTGCACTTGCCTGATCGCAGACGATGAAAGCGGAAGGA 1772

Qy  2136 AAAACTCACACTGACTTCCAAGACACCAACCCTTCTGTCTACAGTTACTTGCCCAGGC 2195
Db  1773 AAAACTCACACTGACTTCCAAGACACCAACCCTTCTGTCTACAGTTACTTGCCCAGGC 1832

Qy  2196 TCTCCAAGCTTCAGCCAATTGGGACCCCCAGCCTCAGAGAGACCTCTGCTCTCTGGCTGCTCAGAGGG 2255
Db  1833 TCTCCAAGCTTCAGCCAATTGGGACCCCCAGCCTCAGAGAGACCTCTGCTCTCTGGCTGCTCAGAGGG 1892
```

```
Db  2256 CATCTGGCAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGGTTAGATGG 2315
Qy  1893 CATCTGGCAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGG 1952

Db  2316 GGCCATCATCTCCACCTTCTTGAAGATGGGTTATTCTTCAAAGGCACCCCATCCCTCTGAG 2375
Qy  1953 GGCCATCATCTCCACCTTCTTGAAGATGGGTTATTCTTCAAAGGCACCCCATCCCTCTGAG 2012

Db  2376 CTACAGCGTTCATTCACCTCTCTGTTTCCAAGAGTTCTTTTGCAGCAAATGTCCTATGTTTGGA 2435
Qy  2013 CTACAGCGTTCATTCACCTCTCTGTTTCCAAGAGTTCTTTTGCAGCAAATGTCCTATGTTTGGA 2072

Db  2436 GGATGAAGGAGGGGGAGGAGGTAAAACATTCTAATTTGCATCATAGGATTTGGAAAAGACGCTAGA 2495
Qy  2073 GGATGAAGGAGGGGGAGGAGGTAAAACATTCTAATTTGCATCATAGGATTTGGAAAAGACGCTAGA 2132

Db  2496 AGCATATGGGAATACATGGCCTGTTTGGGGCATCAACACCACACAGTTTCCTATTTGGGCCCTGTT 2555
Qy  2133 AGCATATGGGAATACATGGCCTGTTTGGGGCATCAACACCACACAGTTTCCTATTTGGGCCCTGTT 2192

Db  2556 AAGTGAATGAGGGGGGAGAGAGAATGGAAGAACATCTTTCACTGCCGCGGCTGCTGGAG 2615
Qy  2193 AAGTGAATGAGGGGGGAGAGAGAATGGAAGAACATCTTTCACTGCCGCGGCTGCTGGAG 2252

Db  2616 GAACCTGATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTCGCACACTCTCTGGAGTC 2675
Qy  2253 GAACCTGATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTCGCACACTCTCTGGAGTC 2312

Db  2676 CCTCCACTGCTTGTACGGAGCTCGGAGACTGGGAACAAAACGTTCCTGCTGACCAAGTGACTGATGGCCCATTT 2735
Qy  2313 CCTCCACTGCTTGTACGGAGCTCGGAGACTGGGAACAAAACGTTCCTGCTGACCAAGTGACTGATGGCCCATTT 2372

Db  2736 CGAAGAAAATGGGCATGTGTGTCGTTAGAAAACAGAACAATGGAGGCTCTTAGTGTGTGCACTTTTCTGCAT 2795
Qy  2373 CGAAGAAAATGGGCATGTGTGTCGTTAGAAAACAGAACAATGGAGGCTCTTAGTGTGTGCACTTTTCTGCAT 2432

Db  2796 TAAATTCAGCGCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGGCAGGCAGCACAGATCAAAC 2855
Qy  2433 TAAATTCAGCGCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGGCAGGCAGCACAGATCAAAC 2492

Db  2856 ATGGAGCCCCCACCCATTGTAGTCCTTGTTCAGGTGAGGCTACAGATGCTGCTTATTGGCA 2915
Qy  2493 ATGGAGCCCCCACCCATTGTAGTCCTTGTTCAGGTGAGGCTACAGATGCTGCTTATTGGCA 2552

Db  2916 GATTCTTCTTCTCCGGTCCTCAAGGTCACCAGGAAACCTTGAAGGAGGCTTGGAACCTTAAGTTGGAAA 2975
Qy  2553 GATTCTTCTTCTCCGGTCCTCAAGGTCACCAGGAAACCTTGAAGGAGGCTTGGAAA 2612

Db  2976 CTCGCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAGGACCCTGAGACGGCCCCTGCCT 3035
Qy  2613 CTCGCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAGGACCCTGAGACGGCCCCTGCCT 2672

Db  3036 CCTGGAGCACCCTGCCGTGGCTGGCTGTGGCCTGACTGCAAGGACCTTGC 3095
Qy  2673 CCTGGAGCACCCTGCGGTGGCTGGCTGTGGCCTGACTGCAAGGACCTTGC 2732

Db  3096 CTTTGGGCTGAGAGCCAACCAGACCCTGACCGAGCTGGAACCTGAGCTTCAATGTGCTCAC 3155
Qy  2733 CTTTGGGCTGAGAGCCAACCAGACCCTGACCGAGCTGGAACCTGAGCTTCAATGTGCTCAT 2792

Db  3156 GGATGCTGGAGCCAAACACCCTTTGCCAGAGACTGAGACTGCAGTTGCAGCTACAGCG 3215
Qy  2793 GGATGCTGGAGCCAAACACCCTTTGCCAGAGACTGAGACTGCAGTTGCAGCTACAGCG 2852

Db  3216 ACTGCAGCTGGTCAGGTGTGCCTCACGTGTGCCTGCAGCTGCCTGCAGGACCTGGCCTGTCTGTGCT 3275
Qy  2853 ACTGCAGCTGGTCAGGTGTGCCTCACGTGTGCCTGCAGCTGCCTGCAGGACCTGGCCTGTGCT 2912

Db  3276 TAGTTGCCAGCCCCAGCCTGAAAGGGGAGCTAGACCTGCAGCAGAACAACCTGGATGACGTTGG 3335
Qy  2913 TAGTTGCCAGCCCCAGCCTGAAAGGGGAGCTAGACCTGCAGCAGAACAACCTGGATGACGTTGG 2972
```

RESULT 6

```
US-09-996-617-5
Sequence 5, Application US/09996617
Patent No. US20020128198A1
GENERAL INFORMATION:
  APPLICANT: Bertin, John
  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
  TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
  FILE REFERENCE: 07334-340001
  CURRENT APPLICATION NUMBER: US/09/996,617
  CURRENT FILING DATE: 2001-11-27
  PRIOR APPLICATION NUMBER: 09/931,071
  PRIOR FILING DATE: 2001-08-15
  PRIOR APPLICATION NUMBER: 09/428,252
  PRIOR FILING DATE: 1999-10-27
  PRIOR APPLICATION NUMBER: 09/340,620
  PRIOR FILING DATE: 1999-06-28
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 5
    LENGTH: 4287
    TYPE: DNA
    ORGANISM: Homo sapiens
US-09-996-617-5

Query Match    78.7%;  Score 4287;  DB 10;  Length 4287;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 4287;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy   523  ATGGCTGGGCGGAGAGCCTGGGCGGCCGCGCCCTGGCCTGGTTACTTGAATTCCTGAAGAAGAAGAGGAG  582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  ATGGCTGGGCGGAGAGCCTGGGCGGCCGCGCCCTGGCCTGGTTACTTGAATTCCTGAAGAAGAAGAGGAG   60

Qy   583  CTGAAGGAGTTCAGCTTCTGCTCGGCCAATAAAGCCACTCCAGAAGCTCTTCGGGGTGAG  642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CTGAAGGAGTTCAGCTTCTGCTCGGCCAATAAAGCCACTCCAGAAGCTCTTCGGGGTGAG  120

Qy   643  ACACCCGCTCAGCCAGAGAGACAGAGTGGCACTGGCCTCGTACCTGGTGGCTCAG  702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  ACACCCGCTCAGCCAGAGAGACAGAGTGGCACTGGCCTCGTACCTGGTGGCTCAG  180

Qy   703  TATGGGGAGCAGCGCGGCCTGGGACTTACACCTGGGACGCAGATGGGGCTGAGG  762
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TATGGGGAGCAGCGCGGCCTGGGACTTACACCTGGGACGCAGATGGGGCTGAGG  240

Qy   763  TCACTGTGCGCCCAAGCCCAGGAAGGGGCAGGAGCCACTCTCCCCTCATTCCCCTACAGCCCA  822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TCACTGTGCGCCCAAGCCCAGGAAGGGGCAGGAGCCACTCTCCCCTCATTCCCCTACAGCCCA  300

Qy   823  AGTGAACCCCACCTGGGGTCTCCCAGCCACCCACCTCCACCGCAGTCGTTGCTAATGCCCTGG  882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AGTGAACCCCACCTGGGGTCTCCCAGCCACCCACCTCCACCGCAGTCGTTGCTAATGCCCTGG  360

Qy   883  ATCCATGAATTGCGGCGGCGGGTCGCACCCAGGGGTTTTGAGACAGCTG  942
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  ATCCATGAATTGCGGCGGCGGGTCGCACCCAGGGGTTTTGAGACAGCTG  420

Qy   943  CCTGACACATCTGGACGCGCGGCTGGAGAGAGAAATCTCTGCCTCCTCCTTCTACCAAGCTCTT  1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CCTGACACATCTGGACGCGCGGCTGGAGAGAGAAATCTCTGCCTCCTCCTTCTACCAAGCTCTT  480

Qy  1003  CCAAGCTCCCCAGACCATGAGTCTTCCAAGCCAGGAGAAATCTCACCCACCA  1062
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CCAAGCTCCCCAGACCATGAGTCTTCCAAGCCAGGAGAAATCTCACCCACCA  540

Qy  1063  GCAGTGCTGGGGAGCTGGGGGATCCCCACCTCAGCCCCAGCTCAGCCCCCAGCAGGAG  1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GCAGTGCTGGGGAGCTGGGGGATCCCCACCTCAGCCCCAGCTCAGCCCCCAGCAGGAG  600

Qy  1123  GCTCCTCTGGGACCCAATTGGCTCTTGGATGAAACGTCAGGAATTACTACAGAAAATCAGA  1182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GCTCCTCTGGGACCCAATTGGCTCTTGGATGAAACGTCAGGAATTACTACAGAAAATCAGA  660

Qy  1183  GAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGGCCCCCATGGGCAGCGGGTGGTAGGAACG  1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGGCCCCCATGGGCAGCGGGTGGTAGGAACG  720

Qy  1243  CCCCACAGGGCGCCACACCAGCCTACAGCCCTACAGCGGGCCCCACCACCACCACCCCCATGGGCAGGCCTTCTGTGTAGGA  1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CCCCACAGGGCGCCACACCAGCCTACAGCCCTACAGCGGGCCCCACCACCACCACCCCCATGGGCAGGCCTTCTGTGTAGGA  780

Qy  1303  GAGAGCCTCGTTCCACATGGCCCCTGGAAAAATGAGGATTTTAACCAAAATTCACACAG  1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GAGAGCCTCGTTCCACATGGCCCCTGGAAAAATGAGGATTTTAACCAAAATTCACACAG  840

Qy  1363  CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTTCAAGAAGCTGG  1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTTCAAGAAGCTGG  900

Qy  1423  CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGAGACTTATTTGCCCA  1482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGAGACTTATTTGCCCA  960

Qy  1483  GGCCTGGGATACCCAAGGAAACCTCGCATAGTCATACTGCAGGGGGGCTTGCTGGAATTGGGAAG  1542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GGCCTGGGATACCCAAGGAAACCTCGCATAGTCATACTGCAGGGGGGCTTGCTGGAATTGGGAAG  1020

Qy  1543  TCAAACACTGGGCCAGGGCCAGGTGAAAGGAAGCCTGGGGGGAGAGGCCAGTTGTGTATGGGGACCGC  1602
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TCAACACTGGCCGCCAGGCAGGTGAAAGGAAGCCTGGGGGGAGAGGCCAGTGTTGTATGGGGACCGC  1080

Qy  1603  TTCCACCATGTCTTCTACTTCAGCTGTCAGCAGAGCTCCCAGTGCCCAGTTCAAGGTGGTGAGTCTC  1662
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTCCACCATGTCTTCTACTTCAGCTGTCAGCAGAGCTCCCAGTGCCCAGTTCAAGGTGGTGAGTCTCT  1140

Qy  1663  GCTGAGCTCATCGGAAAAAGATGGGGACAGCAGCACCACTCCGGGCTCCCATTAGACAGATCCTGTCT  1722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GCTGAGCTCATCGGAAAAAGATGGGGACAGCAGCACCACTCCGGGCTCCCATTAGACAGATCCTGTCT  1200

Qy  1723  AGGCCAGAGAGCGGGCTGCTCTTCATCCTCGATGGTGTGAGATGAGCCAGGATGGGGTCTTGCAG  1782
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AGGCCAGAGAGCGGGCTGCTCTTCATCCTCGATGGTGTGAGATGAGCCAGGATGGGGTCTTGCAG  1260

Qy  1783  GAGCCCAGTTGCTTGTCTGTTCGTGTCGCACTGAGCCACCAGCCGAGAGCCAGGATGCACTGCTG  1842
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GAGCCCAGTTGCTTGTCTGTTCGTGTCGCACTGAGCCACCAGCCGAGAGCCAGGATGCACTGCTG  1320

Qy  1843  GGCAGTTTGCTGTGGGAAAACTATACTTCCCGGAGGCAGTCCTTCCTGATCAGCGGCTGGGACC  1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  GGCAGTTTGCTGTGGGAAAACTATACTTCCCGGAGGCAGTCCTTCCTGATCAGCGGCTGGGACC  1380

Qy  1903  ACAGCTCTGCAGAACCTCCATTCCTTCTTTGGAGCAGCAGGTTGGGTTAGAGGTTCCTGGGG  1962
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  ACAGCTTCTGCAGAACCTCCATTCCTTCTTTGGAGCAGCAGGTTGGGTTAGAGGTTCCTGGGG  1440

Qy  1963  TTCTCTGAGTCCAGCAGGAAGGAAATAATTTCTACAGATATTTCACAGATGAAAGGCAAGCA  2022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  TTCTCTGAGTCCAGCAGGAAGGAAATAATTTCTACAGATATTTCACAGATGAAAGGCAAGCA  1500

Qy  2023  ATTAGAGACCCTTTAGGTTGGTGGTCAAATCAAACAAAGAGAGCTCTGGGGGCCCTGTTGTCC  2082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  ATTAGAGCCTTTAGGTTGGTGGTCAAATCAAACAAAGAGACTCTGGGGCCCTGTTGTCTTGTGCCC  1560

Qy  2083  TGGGTGTCCTGGGCTGGCCTGCGACTGCCTGATGCAGCAGCAGATGAAAGGGAAGGAAAAACTC  2142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TGGGTGTCCTGGGCTGGCCTGCGACTGCCTGATGCAGCAGCAGATGAAAGGGAAGGAAAAACTC  1620

Qy  2143  ACACTGACTTCCAAGACCACCACACCCCTGTTGTCTACATTACCTTGCCCAGGCTCTCCAA  2202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  ACACTGACTTCCAAGACCACCACACCCCTGTTGTCTACATTACCTTGCCCAGGCTCTCCAA  1680

Qy  2203  GCTCAGCCATTGGGACCCCAGCCTCAGAGAGAGACTTCCTCCTCCTCCTCTGGGCTGAGGGCATCTGG  2262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  GCTCAGCCATTGGGACCCCAGCCTCAGAGAGAGACTTCTGCTCTGGGCTGAGGGCATCTGG  1740

Qy  2263  CAAAAAAAGACCCTTTCAGTCCAGATGACCTCAGGAAGCAGCATGGGGGTTAGATGGGGGCCATC  2322
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  CAAAAAAAGACCCTTTCAGTCCAGATGACCTCAGGAAGCAGCATGGGGGTTAGATGGGGGCCATC  ...
```

Db 1741 CAAAAAAAGACCCTTTCAGTCAGATGACCTCAGAGGAGCATGGGTTAGATGGGGCCATC 1800

Qy 2323 ATCTCCACCTTCTTGAACATGGGTATTCTTCTTCAAGAGCACCCATCCCTCTGACCTACAGC 2382
Db 1801 ACTCCACCTTCTTCAAGATGGGTATTCTTCTTCAAGAGCACCCATCCCTCTGAGCTACAGC 1860

Qy 2383 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCACAAGTGTCCTATATGTCTTGGGATGAG 2442
Db 1861 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCACAAGTGTCCTATATGTCTTGGGATGAG 1920

Qy 2443 AAGGGGAGGAGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATAT 2502
Db 1921 AAGGGGAGGAGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATAT 1980

Qy 2503 GGAATACATGGGCCTGTTTGGGGCATCAACCACCACACAGTTCCTATTGGGGCCTGTGTAT 2562
Db 1981 GGAATACATGGGCCTGTTTGGGGCATCAACCACCACACAGTTCCTATTGGGGCCTGTGTAT 2040

Qy 2563 GAGGGGGAGGAGAGATGGAGAACATGGAGAACATCTTTCACTTGCCGCCGCTGTTCAGGGGA 2622
Db 2041 GAGGGGGAGGAGAGATGGAGAACATGGAGAACATCTTTCACTTGCCGCCGCTGTTCAGGGGA 2100

Qy 2623 ATGCAGTGGGTCCCGTCCCTGCCAGCTGCTGCTCGCAGCCACACTCTTCTGGAGTCCTCCAC 2682
Db 2101 ATGCAGTGGGTCCCGTCCCTGCCAGCTGCTGCTCGCAGCCACACTCTTCTGGAGTCCTCCAC 2160

Qy 2683 TGCTTGTTACGGAGACTCGGAACACAAAACGTTCCTGACACAGTGATGGTGGCCCCATTTCGAAGAA 2742
Db 2161 TGCTTGTTACGGAGACTCGGAACACAAAACGTTCCTGACACAGTGATGGTGGCCCCATTTCGAAGAA 2220

Qy 2743 ATGGGCAGTGTGTGTAGAAACAGACATGGAGACTCGCACTTTCTGCATTCATTAAATTC 2802
Db 2221 ATGGGCAGTGTGTGTAGAAACAGACATGGAGACTCGCACTTTCTGCATTCATTAAATTC 2280

Qy 2803 AGCGGCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGGCCACGACAGATCAACATGGAGGC 2862
Db 2281 AGCGGCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGGCCACGACAGATCAACATGGAGGC 2340

Qy 2863 CCCACCACATGGTAGTCCTGTTCAGTCACAGATGATTCTCCCAGTCACAGATGCCATTCGAAGATTCTC 2922
Db 2341 CCCACCACATGGTAGTCCTGTTCAGTCACAGATGATTCTCCCAGTCACAGATGCCATTCGAAGATTCTC 2400

Qy 2923 TTCTCCGTCCTCAAGGTCACCAGAAAACCTGAAGGAGCTGGACTTAAGTCCTAAGTGGAAAACTCGCTG 2982
Db 2401 TTCTCCGTCCTCAAGGTCACCAGAAAACCTGAAGGAGCTGGACTTAAGTCCTAAGTGGAAAACTCGCTG 2460

Qy 2983 AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGGAGCCCCTGCCTCCTCCTGGAG 3042
Db 2461 AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGGAGCCCCTGCCTCCTCCTGGAG 2520

Qy 3043 ACCCTGCGGTTGGGCTGGGCTGTGGGCCCTCACAGATGCGAAGGAGCTGGACTTTTGGG 3102
Db 2521 ACCCTGCGGTTGGGCTGGGCTGTGGGCCCTCACAGATGCGAAGGAGCTGGACTTTTGGG 2580

Qy 3103 CTGAGAGCCAAACCAGACCCTGACCGAGCTGGACTTGAGCTTCAAATGTGCTTCACGGATGCT 3162
Db 2581 CTGAGAGCCAAACCAGACCCTGACCGAGCTGGACTTGAGCTTCAAATGTGCTTCACGGATGCT 2640

Qy 3163 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCGGAGACTGCAAGCTACAGCGACTGCAG 3222
Db 2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCGGAGACTGCAAGCTACAGCGACTGCAG 2700

Qy 3223 CTGGTCAGTGTGGGCTGGGCTGTGGGCCAGGGTGCCAAGGAGACCAGCTGTTTCTGCATTAGTTAGTGCC 3282
Db 2701 CTGGTCAGTGTGGGCTGGGCTGTGGGCCAGGGTGCCAAGGAGACCAGCTGTTTCTGCATTAGTTAGTGCC 2760

Qy 3283 AGCCCCAGCCTTGAAGGAGCTAGACCTGGAGATGAACGGTTGGGCGTGCGA 3342
Db 2761 AGCCCCAGCCTTGAAGGAGCTAGACCTGGAGATGAACGGTTGGGCGTGCGA 2820

Qy 3343 CTGCTCTGTGTGAGGGGGCTCAGGCATCCTGCTGCAAAACTCATACGGCCTGGGGGCTGGACCAG 3402
Db 2821 CTGCTCTGTGTGAGGGGGCTCAGGCATCCTGCTGCAAAACTCATACGGCCTGGGGGCTGGACCAG 2880

Qy 3403 ACAACTCTGAGTGATGAGATGAGGCAGGGAGAACTGAGGGGCCCTGGAGGCAGGGAGGAAACCTCAG 3462
Db 2881 ACAACTCTGAGTGATGAGATGAGGCAGGGAGAACTGAGGGGCCCTGGAGGCAGGGAGGAAACCTCAG 2940

Qy 3463 CTGCTCATCTTCACCGAGACGGAAACCAAATGTAGTGAACCCTACTGAGGCCTGGATACG 3522
Db 2941 CTGCTCATCTTCACCGAGACGGAAACCAAATGTAGTGAACCCTACTGAGGCCTGGATACG 3000

Qy 3523 GGAGGAGATGAGTAAATAGCACATCCTCACTCACAGAGGCAGGGAGACTCGGATCAGAGGACG 3582
Db 3001 GGAGGAGATGAGTAAATAGCACATCCTCACTCACAGAGGCAGGGAGACTCGGATCAGAGGACG 3060

Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGGCTGAGCAAGATCTTCCCAATT 3642
Db 3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGGCTGAGCAAGATCTTCCCAATT 3120

Qy 3643 GCTGAGATTGCAGAGGAAAAGCTCCCCAGAGGGCAGGTAGTACCGGGTGGACTCTTGTGCCGTGCCT 3702
Db 3121 GCTGAGATTGCAGAGGAAAAGCTCCCCAGAGGGCAGGTAGTACCGGGTGGACTCTTGTGCCGTGCCT 3180

Qy 3703 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACATGCAT 3762
Db 3181 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACATGCAT 3240

Qy 3763 GGCCCCACGGGGCCCTGTGGCTACTGGAGTTGACAAGAAAGAACTTGTACCGAGTT 3822
Db 3241 GGCCCCACGGGGCCCTGTGGCTACTGGAGTTGACAAGAAAGAACTTGTACCGAGTT 3300

Qy 3823 CACTTCCCCTGTAGCTGGCCTGGCCTGGAGTTGACAAGAAAGAACTTGATTGAAGA 3882
Db 3301 CACTTCCCCTGTAGCTGGCCTGGCCTGGAGTTGACAAGAAAGAACTTGATTGAAGA 3360

Qy 3883 GAAGCGGGGTGACCGGTTGAGATTGAATTCTGTGTGTGGGGGACCAGTTCCTGGGATCAAC 3942
Db 3361 GAAGCGGGGTGACCGGTTGAGATTGAATTCTGTGTGTGGGGGACCAGTTCCTGGGATCAAC 3420

Qy 3943 CCACAGCACAGCCTGGAGTGGTGGCAGGGCCCTTCGCTGGACATCAAGGCTGAGCCTGGAGCT 4002
Db 3421 CCACAGCACAGCCTGGAGTGGTGGCAGGGCCCTTCGCTGGACATCAAGGCTGAGCCTGGAGCT 3480

Qy 4003 GTGGAAAGCTGTGCACCTCCCTCACTTTGTGGGCTTCCAAGGGGGGGCCCATGGTGGGACATCC 4062
Db 3481 GTGGAAAGCTGTGCACCTCCCTCACTTTGTGGGCTTCCAAGGGGGGGCCCATGGTGGGACATCC 3540

Qy 4063 CTGTTCCAAAATGGCCCACTTTAAAGAGAGAGGGGGAGGGGTGGCTCCTGGAAAGCCAGCCAGGGGTG 4122
Db 3541 CTGTTCCAAAATGGCCCACTTTAAAGAGAGAGGGGGAGGGGTGGCTCCTGGAAAGCCAGCCAGGGGTG 3600

Qy 4123 GAGCTGCATCACATAGTTCTGGAAAACCCCAGCTTCTCCCCGTTCTCCCCTGAAA 4182
Db 3601 GAGCTGCATCACATAGTTCTGGAAAACCCCAGCTTCTCCCCGTTCTCCCCTGAAA 3660

Qy 4183 ATGATCCATAAATGCCCTGGGCCTTCATTCCCGTCACCTCTGTGGTTGCTTTACCACACGC 4242
Db 3661 ATGATCCATAAATGCCCTGGGCCTTCATTCCCGTCACCTCTGTGGTTGCTTTACCACACGC 3720

Qy 4243 GTCCATCCTGAGGAAGTCACCTTCCACCTTCTCACTGATCCCAAGTGACTGCTCCATTCGG 4302
Db 3721 GTCCATCCTGAGGAAGTCACCTTCCACCTTCTCACTGATCCCAAGTGACTGCTCCATTCGG 3780

Qy 4303 AAGGAAACTGGAGCTCTGCTATCGGAAGCCCTGGAGAGACCCAGCTTGTTCTCTGGGAGTTCTAC 4362
Db 3781 AAGGAAACTGGAGCTCTGCTATCGGAAGCCCTGGAGAGACCCAGCTTGTTCTCTGGGAGTTCTAC 3840

Qy 4363 GTTGGCCACTTGGGATCAGGGATCAGGGCTGCAAGTGAAAAGACAAGAAAGATGAGACTCTG 4422
Db 3841 GTTGGCCACTTGGGATCAGGGATCAGGGCTGCAAGTGAAAAGACAAGAAAGATGAGACTCTG 3900

Qy 4423 GTGTGGGGAGGGCCTTGGTGAAAACCAGGAGATCTCATGGCCTGCAACTACTCTGATCCCTCCA 4482
Db 3901 GTGTGGGGAGGGCCTTGGTGAAAACCAGGAGATCTCATGGCCTGCAACTACTCTGATCCCTCCA 3960

RESULT 7
US-09-931-071-5
Sequence 5, Application US/09931071
Patent No. US20020128219A1
GENERAL INFORMATION:
  APPLICANT:
  APPLICANT: Bertin, John
  APPLICANT: Alnemri, Emad S.
  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
  TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
  FILE REFERENCE: 07334-335001
  CURRENT APPLICATION NUMBER: US/09/931,071
  CURRENT FILING DATE: 2002-03-18
  PRIOR APPLICATION NUMBER: 09/428,252
  PRIOR FILING DATE: 1999-10-27
  NUMBER OF SEQ ID NOS: 11
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 5
  LENGTH: 4287
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-931-071-5

Query Match   78.7%;   Score 4287;   DB 10;   Length 4287;
Best Local Similarity   100.0%;   Pred. No. 0;   Mismatches 0;
Matches 4287;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

[Nucleotide sequence alignment between Query (Qy) and Database (Db) follows, spanning positions 1–4287.]

Left column:

```
Qy  1963  TTCTCTGAGTCTCAGCAGAGGAAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCA  2022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  TTCTCTGAGTCTCAGCAGAGGAAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCA  1500

Qy  2023  ATTAGAGCCTTTAGTTGGTCAAATCAAACAAAAGAGCTTCGGGGCCCTGTGTCTTGTGCCC  2082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  ATTAGAGCCTTTAGTTGGTCAAATCAAACAAAAGAGCTTCGGGGCCCTGTGTCTTGTGCCC  1560

Qy  2083  TGGGTGTCCTGGCTGGGCTGGCCTGCACTTGGCACGATGAAGCGGAAGGGAAAAACTC  2142
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TGGGTGTCCTGGCTGGGCTGGCCTGCACTTGGCACGATGAAGCGGAAGGGAAAAACTC  1620

Qy  2143  ACACTGACTTCCAAGACACCACCAACAACCCTCGTCTCTACATTACCTTGCCCAGGCCTCCAA  2202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  ACACTGACTTCCAAGACACCACCAACAACCCTCGTCTCTACATTACCTTGCCCAGGCCTCCAA  1680

Qy  2203  GCTCAGCCAATTGGGACCCCAGCTCAGAGACCCTGCTCCTCGTCTGGGTGCTGAGCCAGGA  2262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  GCTCAGCCAATTGGGACCCCAGCTCAGAGACCCTGCTCCTCGTCTGGGTGCTGAGCCAGGA  1740

Qy  2263  CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAGACGCATGGGTTAGCAGTTTCCAA  2322
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAGACGCATGGGTTAGCAGTTTCCAA  1800

Qy  2323  ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCATCCCTCTGAGCTACACGC  2382
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCATCCCTCTGAGCTACACGC  1860

Qy  2383  TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCACCAATGTCCTAGTGTCTTGGAGGATGAG  2442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCACCAATGTCCTAGTGTCTTGGAGGATGAG  1920

Qy  2443  AAGGGGAGAGGGTAAACATTCTAATTGCATCATAGATTTGGAAAAAGACGCTAGAAGCATAT  2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  GGAATACATGGCCTGTTGGGCATCAACCACAGTTTCCTATTGGGCTGTGTTAAGTGAT  1980

Qy  2503  GGAATACATGGCCTGTTGGGCATCAACCACAGTTTCCTATTGGGCTGTGTTAAGTGAT  2562
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  AAGGGGAGAGGGTAAACATTCTAATTGCATCATAGATTTGGAAAAAGACGCTAGAAGCATAT  2040

Qy  2563  GAGGGGGAGAGAGGATGGAGAACAATCTTTCACTGCCGCGGCTGTGTCTGCTCAGGGGAGGAAACTG  2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GAGGGGGAGAGAGGATGGAGAACAATCTTTCACTGCCGCGGCTGTGTCTGCTCAGGGGAGGAAACTG  2100

Qy  2623  ATGCAGTGGGTTCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGAGTCCTCCAC  2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATGCAGTGGGTTCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGAGTCCTCCAC  2160

Qy  2683  TGCTTGTACGAGACTCGGAACAAAACGTTCCTGACACAAGTGATAGGCCCCATTTCGAAGAA  2742
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TGCTTGTACGAGACTCGGAACAAAACGTTCCTGACACAAGTGATAGGCCCCATTTCGAAGAA  2220

Qy  2743  ATGGGCATGTGTGTAGAAACAACACATGGGAGCTCTTAGTCTGCATTTCTGCATTAAAATTC  2802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  ATGGGCATGTGTGTAGAAACAACACATGGGAGCTCTTAGTCTGCATTTCTGCATTAAAATTC  2280

Qy  2803  AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGCAGCGCACAGATCAACTCAACGGAGC  2862
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGCAGCGCACAGATCAACTCAACGGAGC  2340

Qy  2863  CCCACCCATGGTAGTCTCGTTCAGGTGGGTTCCCAGTCACAGATGCCTAATTGGCAGATTCTC  2922
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  CCCACCCATGGTAGTCTCGTTCAGGTGGGTTCCCAGTCACAGATGCCTAATTGGCAGATTCTC  2400

Qy  2923  TTCTCCGTCCTCAAGGTCACCAGAGAACCTGAAGGAGCTGGACTCACAGATGCTGGAAACTCGCTG  2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  TTCTCCGTCCTCAAGGTCACCAGAGAACCTGAAGGAGCTGGACTCACAGATGCTGGAAACTCGCTG  2460

Qy  2983  AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTGCTCGCTGGAG  3042
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTGCTCGCTGGAG  2520
```

Right column:

```
Qy  3043  ACCCTGCGGTTGGGCTGGGCTGTGGGCCTCACAGCTGATGGAGACTGCAGCTGATTGGG  3102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  ACCCTGCGGTTGGGCTGGGCTGTGGGCCTCACAGCTGATGGAGACTGCAGCTGATTGGG  2580

Qy  3103  CTGAGGAGCCAACCAGAGCCCTGACCAGCTGGACCTGAGCTGAGGACTGCAAATGTCGTCACGGATGCT  3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CTGAGGAGCCAACCAGAGCCCTGACCAGCTGGACCTGAGCTGAGGACTGCAAATGTCGTCACGGATGCT  2640

Qy  3163  GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGGTGCAAACTACAGCGACTGCAG  3222
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGGTGCAAACTACAGCGACTGCAG  2700

Qy  3223  CTGGTCAGCTGTGTGGGCCTCACGTCTGACTGCTGCCAGACCGAGGACCTGGCCTTAGTGCC  3282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CTGGTCAGCTGTGTGGGCCTCACGTCTGACTGCTGCCAGACCGAGGACCTGGCCTTAGTGCC  2760

Qy  3283  AGCCCCAGCCTGAAGGAGCTAGACTGCAGCAGAACAACCTGGAATGACGTTGGCGTTGCCGA  3342
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  AGCCCCAGCCTGAAGGAGCTAGACTGCAGCAGAACAACCTGGAATGACGTTGGCGTTGCCGA  2820

Qy  3343  CTGCTCTGTGAGGGGCTCAGGCATCCTGCCTGCTCCAAACTCATCATACGGGTGACCAG  3402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  CTGCTCTGTGAGGGGCTCAGGCATCCTGCCTGCTCCAAACTCATCATACGGGTGACCAG  2880

Qy  3403  ACAACTCTGAGTGATGAATGAGATTGAGAGGGCAGGGAAACTGAGGGACCAGCAGGAGAAAACCTCAG  3462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  ACAACTCTGAGTGATGAATGAGATTGAGAGGGCAGGGAAACTGAGGGACCAGCAGGAGAAAACCTCAG  2940

Qy  3463  CTGCTCATCTTCAGCAGACGGCGGAAACCAAGTGTGAATGAGACCCCTACTGAGGGGGATACG  3522
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CTGCTCATCTTCAGCAGACGGCGGAAACCAAGTGTGAATGAGACCCCTACTGAGGGGGATACG  3000

Qy  3523  GGAGAGATGAGTAATTAGCACATCTCACTCAAGGCGGCAGAGACTCGGATCAGGAGGCG  3582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  GGAGAGATGAGTAATTAGCACATCTCACTCAAGGCGGCAGAGACTCGGATCAGGAGGCG  3060

Qy  3583  GCTTCCCATGTTGCTCAGGCTAATCTCAAATCTCTGGAGGTGACGTGAGACTCTTCCCAATT  3642
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTGACGTGAGGAACTCTTGTGCGTGCCT  3120

Qy  3643  GCTGAAGATTGCAGAGGAAAGCTCCCCAGAGGTGACGTGAGGAACTCTTGTGCGTGCCT  3702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  GCTTCCCATGTTGCTCAGGCTAATCTCAAATCTCTGGAGGTGACGTGAGACTCTTCCCAATT  3180

Qy  3703  TCTCCTGCCTCTCAAGGGGACCTGCGATACGGAGGGAAGCCTTTGGGGGACTGAACGATGACTTCTGG  3762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  TCTCCTGCCTCTCAAGGGGACCTGCGATACGGAGGGAAGCCTTTGGGGGACTGAACGATGACTTCTGG  3240

Qy  3763  GGCCCCACGCGGGGCCCTGTTGGCCTACTGAGGTAGTTTGACAAAGAAAAAGAAACTTGTACCGAGTT  3822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GGCCCCACGCGGGGCCCTGTTGGCCTACTGAGGTAGTTTGACAAAGAAAAAGAAACTTGTACCGAGTT  3300

Qy  3823  CACTTCCCTGTTAGCTGGGCTCCTACACGGCTGGGGACCAAGCAGCGGGGTCTCTGTGCTTTGGGATGAGA  3882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACTTCCCTGTTAGCTGGGCTCCTACACGGCTGGGGACCAAGCAGCGGGGTCTCTGTGCTTTGGGATGAGA  3360

Qy  3883  GAAGCGGGTGACCGGTTGAGTTGAGATTGAATTCGTGTGTGTGGGGGACCAGTTCCCGGTGAGATCAAC  3942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  GAAGCGGGTGACCGGTTGAGTTGAGATTGAATTCGTGTGTGTGGGGGACCAGTTCCCGGTGAGATCAAC  3420

Qy  3943  CCACAGCACAGCCTGGATGGTGGCAGGGGCCTTCGTGCTTGGACATCAACCAGGCTGAGCT  4002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  CCACAGCACAGCCTGGATGGTGGCAGGGGCCTTCGTGCTTGGACATCAACCAGGCTGAGCT  3480

Qy  4003  GTTGGAAGCTTGTGCACCTCCCTCACTTTGTTGGCTCTCCCAAGGGGGGCCATGTGGACACATCC  4062
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  GTTGGAAGCTTGTGCACCTCCCTCACTTTGTTGGCTCTCCCAAGGGGGGCCATGTGGACACATCC  3540

Qy  4063  CTGTTCCAAAATGGCCCACTTTAAAGGAGGAGGGGATGCTCCTGGAGAAGCCAGCCAGGGGTG  4122
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  CTGTTCCAAAATGGCCCACTTTAAAGGAGGAGGGGATGCTCCTGGAGAAGCCAGCCAGGGGTG  3600

Qy  4123  GAGCTGCATCACATAGTTCTGGAAAAACCCCAGCCTTCTCCCCTTGGGAGTCCTCCTGGAAA  4182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  GAGCTGCATCACATAGTTCTGGAAAAACCCCAGCCTTCTCCCCTTGGGAGTCCTCCTGGAAA  4182
```

RESULT 8
US-09-388-221-1

; Sequence 1, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4422)
US-09-388-221-1

Query Match   76.2%;   Score 4148;   DB 10;   Length 4422;
Best Local Similarity   97.0%;   Pred. No. 0;
Matches 4290;   Conservative 0;   Mismatches 0;   Indels 132;   Gaps 1;

[Nucleotide sequence alignment — Query (Qy) vs Database (Db) — not transcribed in full]

RESULT 9
US-09-388-221-3
Sequence 3, Application US/09388221A
Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regul...
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4197)
US-09-388-221-3

Query Match 75.3%;　Score 4100;　DB 10;　Length 4200;
Best Local Similarity 97.9%;　Pred. No. 0;　Mismatches 0;
Matches 4200; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

```
Qy      523 ATGGCTGGCGGAGCCTGGGCGCCGCCGCCCGTTACTTGGAGTTCCTGAAGAAGAGGAGGAG 582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGCTGGCGGAGCCTGGGCGCCGCCGCCCGTTACTTGGAGTTCCTGAAGAAGAGGAGGAG 60

Qy      583 CTGAAGGAGTTCCAGCTTCTTCGCCGCCAATAAAAGGGGCACACTCCAGGAGGCTCTTCGGGTGAG 642
Db       61 CTGAAGGAGTTCCAGCTTCTTCGCCGCCAATAAAAGGGGCACTCCAGGAGCTCTTCGGGTGAG 120

Qy      643 ACACCCGCTCAGCCACAGAGAAGACGAGTGGGCCATGAAGGAGGTGGGCCTCGTACCTGGTGGCTCAG 702
Db      121 ACACCCGCTCAGCCACAGAGAAGACGAGTGGGCCATGAAGGAGGTGGGCCTCGTACCTGGTGGCTCAG 180

Qy      703 TATGGGGAGCAGCGGGCGCCTGGGACCCTCCACTACCTGGGAGCAGATGGGGCTGAGG 762
Db      181 TATGGGGAGCAGCGGGCGCCTGGGACCCTCCACTACCTGGGAGCAGATGGGGCTGAGG 240

Qy      763 TCACTGTGCGCCCCAAGCCCAGGAAGGGGCAGGCCACTCTCCCTCATTCCCTACAGCCCA 822
Db      241 TCACTGTGCGCCCCAAGCCCAGGAAGGGGCAGGCCACTCTCCCTCATTCCCTACAGCCCA 300

Qy      823 AGTGAACCCCACCTGGGGTCCCCAGCCAGCACCCAGTGCTAATGCCCTGG 882
Db      301 AGTGAACCCCACCTGGGGTCCCCAGCCAACCACCCAGTGCTAATGCCCTGG 360

Qy      883 ATCCATGAATTGCCGCCGGGGTGCACCCAGGGGCTTCAGAGACAGCTG 942
Db      361 ATCCATGAATTGCCGCCGGGGTGCACCCAGGGGTTTTGAGACAGCTG 420

Qy     1003 CCAAGTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAACATCCACA 1062
Db      421 CCTGACACATCTGGACGCCCGCTCGGAGAGAAATTCTCTCGCCTCACTCC 480

Qy     1063 GCAGTGCTGGGGGAGCTGGGGGAATCCCACCCTCCAAGCCACCCAACGCCCCCCCACA 1122
Db      481 CCAAGTCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAACATCCACA 540

Qy     4711 AAGAGATGGGACTCTACCAAGCCCTGAAGGAGAGACCCATCCTCACCTCATTATGGGAACTCTGG 4770
Db      541 GCAGTGCTGGGGGAGCTGGGGGAATCCCACCCTCCAAGCCACCCAGAGCACCAGGAG 600

Qy     1123 GCTCCTGGGACCCAATGGGCCTCTGATGAAGAACTCAGGGAATTACTACACAGAAAATCAGA 1182
Db      601 GCTCCTGGGACCCAATGGGCCTCTGATGAAGAACTCAGGGAATTACTACACAGAAATCAGA 660
```

(...remaining alignment blocks Qy/Db continue with nucleotide sequences through positions 4321–4812 / 3301–4422...)

```
Qy  1183  GAAAGAGAGAGAGAGAAATCAGAGAAAGGCAGGCCCCAGGCCGGGGCCGGGGCTGGGTAGGAACG  1242
          ||||||||||||||||||||||||||||||||||||||||||           ||||||||
Db   661  GAAAGAGAGAGAGAGAAATCAGAGAAAGGCAGGCCCCAGTGGGCAGCGCGGTGGGTAGGAACG   720

Qy  1243  CCCCCACAGGCGCACACCAGCCTACAGCCCCACCCCATGGGCAGCGCGGTGGGTGAGA  1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CCCCCACAGGCGCACACCAGCCTACAGCCCCATGGGGAGCCCTTCTCTGTGAGA   780

Qy  1303  GAGAGGCCTCTGTTCCACATGGGCCCTGGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG  1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GAGAGCCCTGTGTTCCACATGGGCCCTGGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG   840

Qy  1363  CTGCTACTTCTACAAAGACCTCACCCCAGAGAGCCAAGATCCCTGGTCAAGAGAAGCTGG  1422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGCTACTTCTACAAAGACCTCACCCCAGAGAGCCAAGATCCCTGGTCAAGAGAAGCTGG   900

Qy  1423  CCTGATTTATGTGGAGGAGAATCAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCA  1482
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCTGATTTATGTGGAGGAGAATCAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCA   960

Qy  1483  GGCCTGGGATACCCAAGAACCTCGCATAGTCATACTGCAGGGGGGCTGCCTTGGGAATTGGGGAAG  1542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GGCCTGGGATACCCAAGAACCTCGCATAGTCATACTGCAGGGGGCTGCCTTGGAATTGGGGAG  1020

Qy  1543  TCAACACTGGCCAGGCAGGTGAAGGAGAAGCCTGGGGGGAGAGAGCCAGCTGTATGGGGACCGC  1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TCAACACTGGCCAGGCAGTGAAGGAGAAGCCTGGGGAGAGGGCCAGCTGTATGGGGGACCCGC  1080

Qy  1603  TTCCAGCATGTCTTCTACTTCAGTGTGCAGAGAAGGTGGGGTGAGTCTC  1662
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTCCAGCATGTCTTCTACTTCAGTGTGCAGAGAAGGTGGGGGTGAGTCTC  1140

Qy  1663  AGGGCAGAGGGGCTCATCGGAAAAGATGGGGACAGCCACTCCCGGCTCTCCGGCTCCATTAGACAGAAT  1722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GCTGAGCTCATCGGAAAAGATGGGGACAGCCACTCCCGGCTCTCCATTAGACAGAATCCCTGTCT  1200

Qy  1723  AGGCCAGAGACGGGCTGCTCTTCATCCTCGATGGTTGTAGAGATGAGCCACAGAATCCTGGG  1782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AGGCCAGAGACGGGCTGCTCTTCATCCTCGATGGTTGTAGAGATGAGCCACAGAATCCTGTCAG  1260

Qy  1783  GAGCCGAGTTCTGACCTCTGTCTGCACTGGAGCCAGCCGGCGGACAGGACTCAACAGCTG  1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GAGCCGAGTTCTGACCTCTGTCTGCACTGGAGCCAGCCGGCGGACAGGACTCAACATGGCAG  1320

Qy  1843  GGCAGTTTGCTGGGGAAAACTATACTTCCCGAGCCCGGGCCATCTTTCCTGGAAACTCGGCTG  1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  GGCAGTTTGCTGGGGAAAACTATACTTCCCGAGCCATCTTTCCTGATCACGGGCTCGGGACC  1380

Qy  1903  ACAGGCTCTGCAGAACCTCATTCCTTCTTTGGAGCAGGCAGGCACGTTGGGGTAGAGGTTCCTGGGG  1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  ACAGCTCTCAGAACCTCATTCCTTCTTTGTGGAGCAGGCAGGCACGTTGGGGTAGAGGTTCCTGGGG  1440

Qy  1963  TTCTCTGAGTCCAGGAGGAAGCTCATTCTTTCACAGATATATTCACAGATAGAAAAGGCAAGCA  2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  TTCTCTGAGTCCAGCAGGAAGGAATATTTCACAGATATTTCACAGATGAAAAGGCAAGCA  1500

Qy  2023  ATTAGAGCCTTTAGTTGGTTGGTTCAAAATCAAACAAAAGAGCTGGGCCCCTGTGTCTTGTGCCC  2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  ATTAGAGCCTTTAGTTGGTTGGTTCAAATCAAACAAAAGAGCTCTGGGCCCCTGTGTCTTGTGCCC  1560

Qy  2083  TGGGTGTCCTGGCTGGCCTTGCCTCACTTGCCTGATGCAGCAGATGAAGCGGAAGGAAAAACTC  2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TGGGTGTCCTGGCTGGCCTTGCCTCACTTGCCTGATGAAGCAGCAGATGAAGGGAAGGAAAAACTC  1620

Qy  2143  ACACTGACTTCCAAGACCACCACCCCTCTGTTCTACATTACCTTGCCCAGGCTCTCCAA  2202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  ACACTGACTTCCAAGACCACCACCCCTCTGTTCTACATTACCTTGCCCAGGCTCTCCAA  1680

Qy  2203  GCTCAGCCATTGGGGACCCCAGCTCAGAGACCTCTCGCTCTCTGGCTGCCTGCTGCTGAGGGCATCTGG  2262
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  GCTCAGCCATTGGGGACCCCAGCTCAGAGACCTCTCGCTCTCTGGCTTGCTGCTGAGGGCATCTGG  1740


Qy  2263  CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAGCAGCATGGGTTAGAGTTGGGGCCATC  2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  CAAAAAAAGACCCTTTTCAGTCCAGATGGGTAATCCGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC  1800

Qy  2323  ATCTCCACCTTCTTGAAGATAGGGTATTCTTCAAGAGCACCCCATCCCTCCTGTGCTACAGC  2382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ATCTCCACCTTCTTGAAGATAGGGTATTCTTCAAGAGCACCCCATCCCTCCTGTGCTACAGC  1860

Qy  2383  TTCATTCACCTTCTGTTTCCAAGAGTTCTTTGCAGCAAATGCCTATGTCTTGGAGATGAG  2442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TTCATTCACCTTCTGTTTCCAAGAGTTCTTTGCAGCAAATGCCTATGTCTTGGAGATGAG  1920

Qy  2443  AAGGGGGAGGAGGTAAAACATTCTAATTGCATCATCAATAGATTTGGAAAGAAACGGCTAGAAGCATAT  2502
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AAGGGGGAGGAGGTAAAACATTCTAATTGCATCATCAATAGATTTGGAAGAAACGGCTAGAAGCATAT  1980

Qy  2503  GGAATACATGGCCTGTTTGGGGCATCAACAACCACACGTTTCCTATTGGGGCTGTTTAAGTGAT  2562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  GGAATACATGGCCTGTTTGGGGCATCAACAACCACACGTTTCCTATTGGGGCTGTTTAAGTGAT  2040

Qy  2563  GAGGGGGAGAGAGAGATGGAGAACATCTTTCCATCCCGGCTGTCCCATCTCTCGAGTCCTCCAC  2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GAGGGGGAGAGAGAGAGGGAGAACATCTTTCCATCCCGGCTGTCCCATCTCTCGAGTCCTCCAC  2100

Qy  2623  ATGCAGTGGGGTCCCGTTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTCTGGAGAA  2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATGCAGTGGGGTCCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCCAC  2160

Qy  2683  TGCTTGTACGGAGACTCGGAGCAACAAAACGTTCCTGCACAAGTGATGGCCATTTCGAAGAA  2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TGCTTGTACGGAGACTCGGAACAAAACGTTCCTGCACAAGTGATGGCCCATTTCGAAGAA  2220

Qy  2743  ATGGGCATGTGTGTGTTAGAAAACAGACATGAGGCTCTTAGTGTGTGCACTTTTCTGCATTAAATTTC  2802
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  ATGGGCATGTGTGTGTTAGAAAACAGACATGAGGCTCTTAGTGTGTGCACTTTTCTGCATTAAATTTC  2280

Qy  2803  CCCACCATGGTAGTCCTGAAAGAGTTCCAGTGCAGAGGCAGGCACAGATCAAACATGGAGC  2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CCCACCATGGTAGTCCTGAAAGAGTTCCAGTGCAGAGGCAGGCACAGATCAAACATGGAGC  2340

Qy  2863  TTCTCCGTGCCTCCAAGGTCACCAGAAAACCTGAAAGGAGGACTGGAAAGAGAGCTG  2922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTCTCCGTGCCTCCAAGGTCACCAGAAAACCTGAAAGGAGGACTGGAAACTCGGCTG  2400

Qy  2923  AGCCCACACTCTGCAGTGAAAGACTCTTTGTAGAGAAAACCTGAAGACCCCTCGAGACCTCGCTC  2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  AGCCCACACTCTGCAGTGAAAGACTCTTTGTAAGACCCTGAAGACCCCTCGCGTCGTGGAG  2460

Qy  2983  ACCCTGGGCTTGGGCTTGTGGCCTTCACAGCTGCCAAGGAGACCCTGAGGACCCTGAGGACAAG  3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  ACCCTGGGCTTGGGCTTGTGGCCTTCACAGCTGCCAAGGAGACCCTGAGGACCCTGAGGACAAG  2520

Qy  3043  CTGAGAGCCAACCAGACCTTGACCTGGACTGAGCTTCAATGTTGCTAACGTTGCTCACTGCAG  3102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  CTGAGAGCCAACCAGACCTTGACCTGGACTGAGCTTCAATGTTGCTAACGTTGCTCACTGCAG  2580

Qy  3103  GGAGGCCAAACACCCTTTGCCAGAGACTGAGAGCAGCCGAGCTGCAAGCTACAGCGACTGCAG  3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  GGAGGCCAAACACCCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG  2640

Qy  3163  CTGGTCAGCTGTGTGGGCCTCAGCGTCTGACTGCTGCCAGGGACCTGGCCTCTGTTGCTTAGTGCC  3222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  GGAGCCAAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG  2700

Qy  3223  CTGGGTCAGCTGTGTGGGCCTCACGTGTCGCAGGACTGCCTCAGGTGTGCTTTAGTGCC  3282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CTGGGTCAGCTGTGTGGGCCTCACGTGTCGCAGGACTGCCTCAGGTGTGCTTTAGTGCC  2760

Qy  3283  AGCCCCAGCCTGAAGGAGCTAGAGACTGCAGAGAACAACCTGGATGAACGTTGGGGTGCGGA  3342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  AGCCCCAGCCTGAAGGAGCTAGAGACTGCAGAGAACAACCTGGATGAACGTTGGGGTGCGA  2820

Db  3343  CTGCTCTGTGTGAGGGGCTCAGCGCATCCTGCTGCAAACTCATTACGCCTGCTGGGGCTGGACCAG  3402
```

RESULT 10
US-09-388-221-5
; Sequence 5, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4332)
US-09-388-221-5

  Query Match   72.7%; Score 3958; DB 10; Length 4332;
  Best Local Similarity   95.0%; Pred. No. 0;
  Matches 4200; Conservative   0; Mismatches   0; Indels 222; Gaps 2;

[DNA sequence alignment data]

```
Qy   3043 ACCCTGCGGTTGGGCTGTGTGGGCCTCACAGCGACTGAGGACTGCAAGGACCTTGCCTTTGGG 3102
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACCCTGCGGTTGGGCTGTGTGGGCCTTCACACTGCAAGGACACTTTGCCTTTGGG 2580

Qy   3103 CTGAGAGCCAACCAGACCCTGACCGAGCTGAGCTTCAACGGGATGCT 3162
          |||||||||||||||||||||||||||||||||||||||
Db   2581 CTGAGAGCCAACCAGACCCCTGACCGAGCTGAGCTTCAAATGTGCTCACGGGATGCT 2640

Qy   3163 GGAGCCAAACACCTTTGCCAGAGACTGCAGACGACGAGCAG 3222
          |||||||||||||||||||||||||||||||||||
Db   2641 GGAGCCAAAACACCTTTGCCAGAGACTGCAGACAGCGACGACGAGCAG 2700

Qy   3223 CTGGGTCAGCTGTGGCCCTCAGTCTGACTGTCGCAGGGACCTTGACTGTCAGTGCGG 3282
          |||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CTGGGTCAGCTGTGTGGGCCCTCAGCTGACTGTCGCCAGGGACCTTCAGTGCC 2760

Qy   3283 AGCCCCAGCCCTGAAGGAGCTAGACCTGCAGCAGAACAACCTGGATGAGCTGGATGAGCTGCGA 3342
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 AGCCCCAGCCCTGAAGGAGCTAGACCTGCAGCAGAACAACCTGGATGACGTTGGGCGTGGCGA 2820

Qy   3343 CTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCCTGCAAACTCATAACGCCTGAGACCAG 3402
          |||||||||||||||||||||||||||||||
Db   2821 CTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCCTGCAAATCATTACGGCTGG 2869

Qy   3403 ACAACTCTGAGTGATGGAGGAGGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG 3462
          |||||||||||
Db   2870 ---------------------------------------------------------- 2869

Qy   3463 CTGCTCATCTTCAGCAGAGCGGAAAACCAAGTGTGAGATGACCCCTACTGAGGGGCCTGGATACG 3522
          ||||||||||
Db   2870 --------------GGAAACCAAGTGTGATGAGCCCTACTGAGGGGCCTGGATACG 2910

Qy   3523 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGCGCCGCAGAGAGCTTCGGATCGCAGAGAGCGGCG 3582
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2911 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGCGGCAGAGACTTGGATCAGCAGAGGGGCG 2970

Qy   3583 GCTTCCCATGTTGCTCCAGGCTAATCTCAAATCCTTCAAAACTCCTGGACACTTCCCAAATT 3642
          |||||||||||||||||||||||||||||||
Db   2971 GCTTCCCATGTTGCTCAGGCTAATCTCCTGGACGTGACGGTGAGCAAGGATCTTCCCAATT 3030

Qy   3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGGTAGTAGTACCGGTTGGGAACTCTTGTGTCGCTGCCT 3702
          |||||||||||||||||||||||||||||||||||||||
Db   3031 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGGTAGTACCGGTTGGGAACTCTTGTGCCTGCCT 3090

Qy   3703 TCTCCTGCCTCTCAAGGGGACCTGCATACCGAGCCCTTTGGGGTGTTGGGTGAGGACTTCTGG 3762
          ||||||||||||||||||||||||||||||||||||||||
Db   3091 TCTCCTGCCTCTCAAGGGGACACTGCATACGGAGGCTGAGGGTGACGATGACTTCTGG 3150

Qy   3763 GGCCCCCACAGGGGGCCTTGGCCTACTGAGGGTAGTTGACAAAGAAAAGAAACTTGTACCGGAGTT 3822
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   3151 GGCCCCCACAGGGGGCCCTGTGGGCTACTGAGGGTAGTTGACAAAGAAAAGAAACTTGTACCGGAGTT 3210

Qy   3823 CACTTCCCTGTAGCTGGCTCCTACGCCTGGCCCAACACGCGGGTGTCTCTGCTTTGTGATGAGA 3882
          |||||||||||||||||||||||||||||||||||
Db   3211 CACTTCCCTGTAGCTGGCTCCTACGCCTGGCCCAACACGCGGGTCTCTGCTTTGTGTGATGAGA 3270

Qy   3883 GAAGCCGGTGACCGGTTGAGATTGAATTCTGTGTGTGGGAGGACCCAGTTCCTGGGTGAGATCAAC 3942
          ||||||||||||||||||||||||||||||||||||
Db   3271 GAAGGCGGTGACCGGTTGAGATTGAATTCTGTGTGTGGGAGGACCAGTTCCTGGGTGAGATCAAC 3330

Qy   3943 CCACAGCACAGCTGGATGGTGGCAGGGGCCTTCGTGTGGGACATCAAGGCCTGAGCGCTGGAGCT 4002
          |||||||||||||||||||||||||||||||||||
Db   3331 CCACAGCACAGCTGGATGGTGGCAGGGGCCTTCTGCTCTGGACATCAAGGCTGAGCCTGGAGCT 3390

Qy   4003 GTGGAAGCTGTGCACCTCCCTCACTTTGTTGGCTCTTCCAAGGGGGCCATGTGGACACATCC 4062
          |||||||||||||||||||||||||||||||||
Db   3391 GTGGAAGCTGTGCACCTCCCTCACTTTGTTGGCTCTTCCAAGGGGGCCATGTGGACACATCC 3450

Qy   4063 CTGTTCCAAAATGGCCCCACTTTAAAGAGGAGGGGGATGCTCCTGGAGAAAGCCAGCCAGGGTG 4122
          ||||||||||||||||||||||||||||||||||
Db   3451 CTGTTCCAAAATGGCCCCACTTTAAAGAGGAGGGGATGCTCCTCTGGAGAAAGCCAGCCAGGGTG 3510

Qy   4123 GAGCTGCATCACATAGTTCTTGAAAACCCAGCTTCTTCCCCCTTGGGGAGTTCCTCCTGAAA 4182
          ||||||||||||||||||||||||||||||||||||
Db   3511 GAGCTGCATCACATAGTTCTTGGAAAACCCAGCTTCTTCCCCCTTGGGGAGTTCCTCCTGAAA 3570

Qy   4183 ATGATCCATAATGACCTTGCGCGCTTCATTCCCGTCACCTCGTTGTTGGGTGGTTGGTTACCACGC 4242
          ||||||||||||||||||||||||||||||||||||
Db   3571 ATGATCCATAATGCCCTTGCGCTTCATTCCCGTCACCTTCTTGGTGGTGTTGGTTTACCACGGC 3630

Qy   4243 GTCCATCCTGAGGGAAGTCACCTTCCACCTCTACCTGATCCCAAAGTGACTGCTCCATTCGG 4302
          |||||||||||||||||||||||||||||||||||
Db   3631 GTCCATCCTGAGGGAAGTCACCTTCCACCTCTACCTGGATCCCAAAGTGACTGCTCCATTCGG 3690

Qy   4303 ---------------------------------------------------------- 4302
Db   3691 AAGGCCATAGATGAGTGAGACTAGAAATGAATTCAGTTGTGCGAAATCACAAGCCACCCCG 3750

Qy   4303 CTGACCCCACTTTATTATGGGCTGTTGTTACTGTCTTGGGGTTCTGGGTCTCGGTTCGGGGATGCTG 4302
          |||||||||||||||||||||||||||||||||||||
Db   3751 CTGACCCCACTTTATTATGGGCTGTTGTTACTGTCTTGGGGTTCTGGGTCTCGGTTCAGGGATGCTG 3810

Qy   4303 --------------AAGGAAACTGAGCTCTGCTTATCGAGGCTGCGTTC 4350
          ||||||||||||
Db   3811 GAAATACTCCCCAAGGAAACTGGAGCTTCTGCTTATCGAGGCCGTTC 3870

Qy   4351 TCGGAGTTCTACGTTGGCCCACTTGGGATCAGGGATCAGGCGTGTCAAGTGAAAGACACAGAAA 4410
          ||||||||||||||||||||||||||||||||||
Db   3871 TCGGAGTTCTACGTTGGCCCACTTGGGATCAGGGATCAGGCGTGCAAGTGAAAGACACAAGAAA 3930

Qy   4411 GATGAGACTCTGGGTGTGGGAGGCCTTGGGTGAAACCAGGAGATCTCATGCCTGCCAACTACT 4470
          ||||||||||||||||||||||||||||||||||
Db   3931 GATGAGACTCTGGGTGTGGGAGGCCTTGGTGAAACCAGGAGAATCTCATGCCTGCCAACTACT 3990

Qy   4471 CTGATCCCTCCAGCCCGCCATAGCCGTACCTTCACCCTCTGGATGCCGCGCCAGTTGCTGCAC 4530
          ||||||||||||||||||||||||||||||||||
Db   3991 CTGATCCCTCCAGCCCGCCATAGCCGTACCTTCACCCTCTGGATGACGCCAGTTGCTGCAC 4050

Qy   4531 TTTGTGGACCAGTCATCCGAGACGACAGCTGATTAGCCCGAGTGACATCGGTGGAGGGGTTGTCTTG 4590
          |||||||||||||||||||||||||||||||||||
Db   4051 TTTGTGGACCAGTCATCCGAGACGACAGCTGATTAGCCCGAGTGACATCGGTGGAGGTTGTCTTG 4110

Qy   4591 GACAAACTGCATGGACAGGGTGCTGAGCCAGGAGCAGCTACGAGGAGGGTTGCTGGCTGAGGAAC 4650
          ||||||||||||||||||||||||||||||||||
Db   4111 GACAAACTGCATGGACAGGGTGCTGACCGAGGAGCAGCTACGAGGGTTGCTGGCTGAGAAC 4170

Qy   4651 ACGAGGCCCAGCCAGATGCGGAAAGCTGTTCAGCTTGAGCCAGTCTCTGGGACCGGGAAGTGC 4710
          ||||||||||||||||||||||||||||||||||
Db   4171 ACGAGGCCCAGCCAGATGCGGAAGCTGTTCAGCTTGAGCCAGTCTCTGGGACCGGAAGTGC 4230

Qy   4711 AAAGATGGACTCTACCAAGCCCTGAAGGAGCAGCCATCCTCACCTCATTATGTGAACTCTGG 4770
          |||||||||||||||||||||||||||||||||||
Db   4231 AAAGATGGACTCTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATTATGGAACTCTGG 4290

Qy   4771 GAGAAGGGGCAGCAAAAAAGGGGGAGACTCCTGCCCACTCACAGCTGA 4812
          |||||||||||||||||||||||||||||||||||
Db   4291 GAGAAGGGGCACGCAAAAAGGGGACTCCTGCCCACTCAGCAGCTGA 4332
```

RESULT 11
US-09-388-221-9
; Sequence 9, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1e1 Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;  FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)..(4362)
;    OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;    OTHER INFORMATION: Construct
US-09-388-221-9

Query Match      62.5%;   Score 3400.4;   DB 10;   Length 4556;
Best Local Similarity   93.7%;   Pred. No. 0;   Mismatches 236;
Matches 3555;   Conservative 0;   Mismatches 236;   Indels 3;   Gaps 1;

Qy    523 ATGGCTGGCGCGGAGCCCTGGGCGCCCCGGCCGCCCGTTACTTGGAGTTCCTGAAGAAGGAGGAG  582
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGCGCGGAGCCCTGGGCGCCCCGGCCGCCCGTTACTTGGAGTTCCTGAAGAAGGAGGAG   60

Qy    583 CTGGAAGGAGTTCCAGCTTCTTGCTCGCCAATAAAGCGCACTCCAGGAGCTTCTTCGGGGTGAG  642
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGGAAGGAGTTCCAGCTTCTTGCTCGCCAATAAAGCGCACTCCAGGAGCTTCTTCGGGGTGAG  120

Qy    643 ACACCCGCTCAGCCAGAGAGACGAGTGGCCATGGAGGTGGCCTGTACCTGGTGGCTCAG      702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACACCCGCTCAGCCAGAGAGACGAGTGGCCTTGGTACCTCCGTACCTGGTGGCTCAG      180

Qy    703 TATGGGGAGCAGCACGGGGGCCTGGGAGCCCCTCCCATACCTGGGGAGCAGATGGGGGCTGAGG  762
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATGGGGAGCAGCACGGGGGCCTGGGAGCCCCTCCCATACCTGGGGAGCAGATGGGGGCTGAGG  240

Qy    763 TCACTGTGCGCCCAAGCCCACTCTCCCTCCCCATTCCCCACCCAGCCCA             822
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTGTGCGCCCAAGCCCACTCTCCCTCCCCATTCCCCACCCAGCCCA             300

Qy    823 AGTGAACCCCACCTGGGGTCTCCCAGCCAACCCACCTCACCGCAGTGCTAATGCCCCTGG      882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGTGAACCCCACCTGGGGTCTCCCAGCCAACCCACCTCACCGCAGTGCTAATGCCCCTGG      360

Qy    883 ATCCATGAAATTGCCGGACGCGCGGCGGGGGGGCGGAAATCTGCCTCACTGTTTGAGACAGCTG  942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCCATGAAATTGCCGGACGCGCGGCGGGGGGGCGGAAATCTGCCTCACTGTTTGAGACAGCTG  420

Qy    943 CCTGACACATCTGGACGCGCGCTGGAGAGAAATCTCTGCCTCACTCCTCCTACCAAGCTCTT  1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTGACACATCTGGACGCGCGCTGGAGAGAAATCTCTGCCTCACTCCTCCTACCAAGCTCTT   480

Qy   1003 CCAAGCTCCCAGACCATGAGTCTCCAAGCCACCAGGAGTCACCCAAGCGCCCCCCACATCCACA  1062
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCAAGCTCCCAGACCATGAGTCTCCAAGCCACCAGGAGTCACCCAAGCGCCCCCCACATCCACA  540

Qy   1063 GCAGTGCTGTGGGGAGCTGGGGATCCCCAGCCTCACCTACCCAGCAGAGAGCAGCAGAG      1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCAGTGCTGTGGGGAGCTGGGGATCCCCAGCCTCACCTACCCAGCAGAGAGCAGCAGAG      600

Qy   1123 GCTCCTGGGACCCCAAATGGCCTCTGGATGAAAACGTCAGGAATTTACTCACAGAGAAATCAGA  1182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCTGGGACCCCAAATGGCCTCTGGATGAAAACGTCAGGAATTTACTCACAGAGAAATCAGA  660

Qy   1183 GAAAGAGAGAGAGAGAGAAATCAGAGAGAAAGGCAGCAGGCCCCATGGGCAGCAGGGTGGAAGC  1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAAAGAGAGAGAGAGAGAAATCAGAGAGAAAGGCAGCAGGCCCCATGGGCAGCAGGGTGGAAGC  720

Qy   1243 CCCCCACAGGGCGCACCAGCCTACAGCCCCACCACCCCCCATCCCTTACGCCCTTCTGTGAGA  1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCCCACAGGGCGCACCAGCCTACAGCCCCACCACCCCCCATCCCTTACGCCCTTCTGTGAGA   780

Qy   1303 GAGAGGCCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG  1362
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAGGCCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG   840

Qy   1363 CTTGCTACTTCTACAGAAGAACCTCACCCCAGAAGCCCAAGATCCCTGGTCAAGAGAAGCTGG  1422
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTTGCTACTTCTACAGAAGAACCTCACCCCAGAAGCCCAAGATCCCTGGTCAAGAGAAGCTGG  900
```

```
Qy   2563 GAGGGGGAGAGAGAGGAGGATGGAGAGAACATCTTTCACTCGCCGGCTGTTCAGGGGAGGAACCTG 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GAGGGGGAGAGAGAGGAGGATGGAGAGAACATCTTTCACTCGCCGGCTGTTCAGGGGAGGAACCTG 2100

Qy   2623 ATGCAGTGGGGTCCCGTCCCTGCAGTGCTGCAGCCACACTCTCTTGGAGTCCCTCCAC 2682
          
Db   2101 ATGCAGTGGGGTCCCGTCCCTGCAGTGCTGCAGCCACACTCTCTTGGAGTCCCTCCAC 2160

Qy   2683 TGCTTGTACGAGACTCGGGACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAA 2742
          
Db   2161 TGCTTGTACGAGACTCGGGACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAA 2220

Qy   2743 ATGGGGCATGTGTGTAGAAACAGACATGGAGGCTCTTAGTTGTCACTTTCTGCATTAAATTC 2802
          
Db   2221 ATGGGGCATGTGTGTAGAAACAGACATGGAGGCTCTTAGTTGTCACTTTCTGCATTAAATTC 2280

Qy   2803 AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGAGCAGGCAGCAGCACAGATCAAACATGGAGC 2862
          
Db   2281 AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGAGCAGGCAGCAGCACAGATCAAACATGGAGC 2340

Qy   2863 CCCAACCATGGTAGTCCCTGTTCAGGTGGGTCCCAGTCACACAGATGCCTATTGGCAGATTCTC 2922
          
Db   2341 CCCAACCATGGTAGTCCCTGTTCAGGTGGGTCCCAGTCACACAGATGCCTATTGGCAGATTCTC 2400

Qy   2923 TTCTCCGGTCCTCAAGGTCACCAGAAACCTGAAGGAGAGTCGGACCTAAGTGGAAAACTCGCTG 2982
          
Db   2401 TTCTCCGGTCCTCAAGGTCACCAGAAACCTGAAGGAGAGTCGGACCTAAGTGGAAAACTCGCTG 2460

Qy   2983 AGCCACTCTCGCAGTGAAGAGTCTTTGTAAGACCCTGAGAGCGCCCTCGCTGCCTCCTGGAG 3042
          
Db   2461 AGCCACTCTCGCAGTGAAGAGTCTTTGTAAGACCCTGAGAGCGCCCTCGCTGCCTCCTGGAG 2520

Qy   3043 CTGAGGAGCCAACCAGACCCTGACCGAGCTGAGGACTGCAGCTGAACTGCCTTGCCTTTGGG 3102
          
Db   2521 ACCCTGCGGTTGGCTGGCTGGCTCACAGCTGAGGACTGCAGCTGAACTGCCTTGCCTTTGGG 2580

Qy   3103 CTGAGGAGCCAACCAGACCCTGACCGAGCTGAGGACTGCAGCTGAACTGCCTTGCCTTTGGG 3162
          
Db   2581 CTGAGGAGCCAACCAGACCCTGACCGAGCTGAGGACTGCAGCTGAACTGCCTTGCCTTTGGG 2640

Qy   3163 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGGAGTGCAAACTACAGCGACTGCAG 3222
          
Db   2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGGAGTGCAAACTACAGCGACTGCAG 2700

Qy   3223 CTGGTCAGCTGTGGCCTCAGTCTGACTGCTGCCAGGACCTGGCCTCTGTGTGCTTAGTGCC 3282
          
Db   2701 CTGGTCAGCTGTGGCCTCAGTCTGACTGCTGCCAGGACCTGGCCTCTGTGTGCTTAGTGCC 2760

Qy   3283 AGCCCCAGCCTGAAGGAGCTAGACCTGCAGAGAACAACCTGGATAACGTTGGCGGTGCGA 3342
          
Db   2761 AGCCCCAGCCTGAAGGAGCTAGACCTGCAGAGAACAACCTGGATGACGTTGGGCGGTGCGA 2820

Qy   3343 CTGCTCATCTTCAGCAGACCGGAAACCAAGTGTGATACGACCCCTACTGAGGGGCTGGATACG 3402
          
Db   2821 CTGCTCGTGTGAGGGGCTCAGGCATCCTGCCTGCAAACTCATACGCGTGGGGCTGGACCAG 2880

Qy   3403 ACAACTCTGAGTGATGATGAGAGGATGAGGGCCCTGGAGCGAGGAGAAACCTCAG 3462
          
Db   2881 ACAACTCTGAGTGATGATGAGAGGATGAGGGCCCTGGAGCGAGGAGAAACCTCAG 2940

Qy   3463 CTGCTCATCTTCAGCAGACCGGAAACCAAGTGTGATATGACCCCTACTGAGGGGCTGGATACG 3522
          
Db   2941 CTGCTCGTGTGAGGGGCTCAGGCATCCATACGCGTGGGGCTGGGGCTGGACCAG 3000

Qy   3523 GGAGAGATGAGTAATTAGCACATCCCTCCACTCAAGCGCGCAGAGACTCGGATGACGCCTGGATGCG 3582
          
Db   3001 GGAGAGATGAGTAATTAGCACATCCCTCCACTCAAGCGCGCAGAGAGGGGCG 3060

Qy   3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTGGGACGTGAGCGAAGATCTTTCCCAATT 3642
          
Db   3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTGGGACGTGAGCGAAGATCTTTCCAATT 3120

Qy   3643 GCTGAGATTGCAGAGAGAAAAGCTCCCAGAGGTAGTACCGGTGGAACTTCTTGTGGGTGCCT 3702
          
Db   3121 GCTGAGATTGCAGAGAGAAAAGCTCCCAGAGGTAGTACCGGTGGAACTTCTTGTGGGTGCCT 3180

Qy   3703 TCTCCTGCCTTCTCAAGGGGACCTGCATACAGAAGCCTTTGGGGACTGAGGAGACTTCTGG 3762
          
Db   3181 TCTCCTGCCTTCTCAAGGGGACCTGCATACAGAAGCCTTTGGGGACTGAGGAGACTTCTG 3240

Qy   3763 GGCCCCACGGGGGCCCTTGGGCTACTGAGGGTAGTTGACAAAAGAAAAGAACTTGTTGTACCGAGTT 3822
          
Db   3241 GGGCCCTGAAGGAAAATGTGGATGTGAGTGATTGAGTAGAGGACAAAACAGATACAGCGGTT 3300

Qy   3823 CACTTCCCTGTTAGCTCGGCTTCCTACCGCCTGGGCCCCAACACGGGGTCTCTTGCTTTGTGATGAGA 3882
          
Db   3301 TGGTTCCCCACTGCTCCTCGGCTGTATCGTGCTGCGGTCAGCCACACAGGCGCCTTCCTGTGGTAAGG 3360

Qy   3883 GAAGCGGGTGACCGTTGAGATTGAATTCGTGTGTGGTGGGGACCAGTTCCTTCTGG---GTGAGATC 3939
          
Db   3361 GATGAGGTCACAGTGACAGATTGCGTTTGGTTCCTGGAGTCAGCCACCTTGAAGGCTGGGACCTG 3420

Qy   3940 AACCCACAGCACAGCTGGATGGTGGGCAGGGGGCCTCTGCTGACCATCAAGGCTGAAGTGA 3999
          
Db   3421 CAGCACCATGAACAGTGGCTGCTGGGCGGGGCCCCTTGTTGTATCGCTACTGCAGCAGAG 3480

Qy   4000 GCTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGGTCTCCAAGGGGGGCCATGTGGACACA 4059
          
Db   3481 GAGGCTGTGCGCGAGAAATCCACCTCCCCACTTCATCTCCCCAGAGGTGGTGGGAGGTC 3540

Qy   4060 TCCCTGTTCCAAATGGCCCACTTTAAGAGAGGAGGGGGATGCTCTGGAGGAGGCCAGCAGG 4119
          
Db   3541 TCCTGGTTTCCGTTGCCCATTTTAAGGAATGAAGGGATGGTCCTGGAGCATCCAGCCCGG 3600

Qy   4120 GTGGGACCTGCATCACATAGTTCTGGAGAAAACCCCAGCTTTGCCCCCTTGGGGAGTCCTCCTG 4179
          
Db   3601 GTGGGAGGCCCTTCTATGCTGTCCTGGGAAAGCCCCAGCTTTGCCTGGTGTTGCTTTACCAC 3660

Qy   4180 AAAATGATCCATAAATGCCCTGCGCTTCATTCCCGTCACCCTGTGGTTGCTTTACCAC 4239
          
Db   3661 CGGATCGGCCAGTGGGACTCGCCTCTCCCATCCCATCCTCCTCCTCCTCGATGGGGCATCCTCGCTG 3720

Qy   4240 CGCGTCCATCCTGAGGAAGTCACCTTCCACCTCACCTGATCCGAAGTGACTGACTTCCAATT 4299
          
Db   3721 CACCCCACCCCGAAGATATTAAGTTCCACTTGTGTCCCCAGGCAGGACGCCTTGCTA 3780

Qy   4300 CGGAAGGAACTGGA 4313
          
Db   3781 ACAAAGGCGGATAGA 3794
```

Query Match   59.0%; Score 3210.4; DB 10; Length 4466;

Best Local Similarity 91.3%; Conservative 0; Pred. No. 0; Mismatches 236; Indels 93; Gaps 2;
Matches 3465;

RESULT 13
US-09-895-298-22
; Sequence 22, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: PZ035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29950
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/112,809
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-298-22

Query Match   47.3%; Score 2572.6; DB 11; Length 2657;
Best Local Similarity   99.4%;
Matches 2614; Conservative   0; Mismatches 14; Indels 3; Gaps 3;

[DNA sequence alignment — Qy (query) vs Db (database) rows with match bars, positions 2161–4313]

```
RESULT 14
US-09-956-712-10

; Sequence 10, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT: Susan M. Freier
;  TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
;  FILE REFERENCE: RTS-0326
;  CURRENT APPLICATION NUMBER: US/09/956,712
;  CURRENT FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 10
;  LENGTH: 96649
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)...(96649)
;  OTHER INFORMATION: n = A,T,C or G
US-09-956-712-10

Query Match      31.4%;  Score 1710.8;  DB 11;  Length 96649;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1712;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;
```

(Nucleotide sequence alignment — Qy = Query, Db = Database — coordinate numbers shown at the start and end of each line; vertical bars denote matched bases.)

Left column:

```
Qy    5034 TGGAGCTGGCAAGACCCCTCAGACCTCAGTAGAGCGCAGCGCCAA 5093
           ||||||||||||||||||||||||||||||||||||||||||||
Db    2226 TGGAGCTGGCAAGACCCCTGCAGACCTCATCAGTAGAGCGCAGCGCA 2285

Qy    5094 GCCTAGAGCGCCTCCCGGATCCCATCCCAGGCGGCAAGGAGGGAATAGAGCCTCATCTGGTGGGCGGCAGCAACCAT 5153
           ||||||||||||||||||||||||||||||||||||||||||||
Db    2286 AGCCTAGAGCGCCTCCCGGATCCCATCCCAGGCGGCAAGGAGGGAATAGAGGGGCGGCAGCAACCAT 2345

Qy    5154 TTGCCTCTGGCTGTGTCACAGGGGTGAGCCCACGG 5213
           ||||||||||||||||||||||||||||||||||||||||
Db    2346 TTGCCTCTGGCTGTGTCACAGGGGTGAGCCGTTCAGCGTGGGCAGGGAGGCCACG 2405

Qy    5214 TGGATTCTTGGCTTTGTACAGGAGACATCTACAGAGACAGAGTAAAGTGGAAG 5273
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    2406 TGGATTCTTGGCTTTGTACAGGAGACATCTACAGAGACAGAGTAAAGTGGAAG 2465

Qy    5274 GAAGTTTATTCAGAAAAATAAAGAGGAGTATCACAGCTTCTTTTAGAATTGTTCTGACAGGCTT 5333
           |||||||||||||||||||||||||||||||||||||||||
Db    2466 GAAGTTTATTCAGAAAAATAAAGAGGAGTATCACAGCTTCTTTTAGAATTGTTCTGACAGGCTT 2525

Qy    5334 TCCAGTTTTTACAGAAAAACCCCTATAAATTAAAATTTTTTTTTTTTTTTACTTAAATTTAAGAATTA 5393
           ||||||||||||||||||||||||||||||||||||||||
Db    2526 TCCAGTTTTTACAGAAAAACCCCTATAAATTAAAATTTTTTTTTTTTTACTTAAATTTAAGAATTA 2585

Qy    5394 AAAAAATACAAAAAAGAAAAAAATGAAAAATTAAAGGAATAAGAAGTTACCTAC 5444
           |||||||||||||||||||||||||||||||||||||||||
Db    2586 AAAAAATACAAAAAAGAAAAAAATGAAAAATAAGGAATAAGAAGTTACCTAC 2636
```

```
Qy    1168 TACACAGAAAATCAGAGAAGAGAGAGAGAGAGAGAAATCAGAGAGAGAAAGAGAGAGGGAGGGCCCCCATGGGCA 1227
           ||||||||||||||||||||||||||||||||||||||||
Db   30558 TACTTAGAAATCAGAGAAAAGAGAGAGAAATCAGAGAGAAAGGCAGGGCCCCCATGGGCA 30617

Qy    1228 GCGGTGGTAGGAAGCGCCCCACAGGCGCACCACGCCTACAGGGTAGCCCCCATGGGGC 1287
           ||||||||||||||||||||||||||||||||||||||||
Db   30618 GCCGGTGGTAGGAAGCGCCCCACAGGCGCACCACGCCTACAGGGTAGCCCCCATGGGGC 30677

Qy    1288 GAGCCTTCTGTGAGGAGAACGGCCCTCGTTCCACATGGCCATCAGCCGCCTACACCACCACCCATGGG 1347
           ||||||||||||||||||||||||||||||||||||||||
Db   30678 GAGCCTTCTGTGAGGAGAACGGCCCTCGTTCCACATGGCCATCAGCCGCCTACACCACCACCCATGGG 30737

Qy    1348 CAAAAAATTCACACAGCTGCCTACTTCTACAAAGACCTCACCCAGAAAAATGAGGATTTTAAC 1407
           ||||||||||||||||||||||||||||||||||||||||
Db   30738 CAAAAAATTCACACAGCTGCCTACTTCTACAAAGACCTCACCCAGAAAAATGAGGATTTTAAC 30797

Qy    1408 GTCAAGAGAAGCTGGCCTGATTATGTGGAGGAGAATCGAGAGACATTTAATTGAGATCAGA 1467
           ||||||||||||||||||||||||||||||||||||||||
Db   31818 GTCAAGAGAAGCTGGCCTGATTATGTGGAGGAGAATCGAGAGACATTTAATTGAGATCAGA 31877
```

Right column:

```
Db   30798 GTCAAGAGAAGCTGGCCTGATTATGTGGAGGAGAATCGAGAGACATTTAATTGAGATCAGA 30857

Db    1468 GACTTATTTGGCCCAGGCCTGGATACCCAAGAGAACTCGCATAGTCATACTGCAGGGGCT 1527

Db   30858 GACTTATTTGGCCCAGGCCTGGATACCCAAGAGAACTCGCATAGTCATACTGCAGGGGCT 30917

Qy    1528 GCTGGAATTGGGAAGTCAACACTGGCCAGGCAGGTGAAAGAAGCCTTGGGGGGAGAGGCCAG 1587

Db   30918 GCTGGAATTGGGAAGTCAACACTGGCCAGGCAGGTGAAAGAAGCCTTGGGGGGAGAGGCCAG 30977

Db    1588 CTGTATGGGGAACCGCTTCCAGCATGTCTTCTTACTTCAGCTGCCAGGAGGCTGGCCCAGTCC 1647

Db   30978 CTGTATGGGGAACCGCTTCCAGCATGTCTTCTTACTTCAGCTGCCAGGAGGCTGGCCCAGTCC 31037

Qy    1648 AAGGTGGGTGAGTCTCGCTGAGCTCATCGGAGAAAGATGGGACAGCAGCCACTCCGGCTCCATT 1707

Db   31038 AAGGTGGGTGAGTCTCGCTGAGCTCATCGGAGAAAGATGGGACAGCAGCCACTCCGGCTCCATT 31097

Qy    1708 AGACAGATCCGTCTTCTAGGGCCAGAGCGCGGGCTGTCTCTTCATCCTCGATGGTGTAGGATGAGCCA 1767

Db   31098 AGACAGATCCGTCTTCTAGGGCCAGAGCGCGGGCTGTCTCTTCATCCTCGATGGTGTAGGATGAGCCA 31157

Qy    1768 GGATGGGTCTTGCAGGAGCCGAGTTCTGACGTCTGCACTGAGGACGCCAGCCACCAGCCG 1827

Db   31158 GGATGGGTCTTGCAGGAGCCGAGTTCTGACGTCTGCACTGAGGACGCCAGCCACCAGCCG 31217

Qy    1828 GCGGATGCACTGCTGCTGGGGCAGTTTGCTGGGGGAAAAACTATACTTCCCGAGGCATCCTTCCTG 1887

Db   31218 GCGGATGCACTGCTGCTGGGGCAGTTTGCTGGGGGAAAAACTATACTTCCCGAGGCATCCTTCCTG 31277

Qy    1888 ATCACGCGGCTCGGACCACAGCTTCTGCAGAACCTCATTCCTTCCTTCGTTTGGAGCACCAGTTGG 1947

Db   31278 ATCACGCGGCTCGGACCACAGCTTCTGCAGAACCTCATTCCTTCCTTCGTTTGGAGCACCAGTTGG 31337

Qy    1948 GTAGAGGTCCTGGGGTTCTCTGAGTCCAGCAGGAAGGAGAATATTTCTACAGATAATTTCACA 2007

Db   31338 GTAGAGGTCCTGGGGTTCTCTGAGTCCAGCAGGAAGGAGAATATTTCTACAGATAATTTCACA 31397

Qy    2008 GATGAAAAGGCAAGCAAATTAGAGCCCTTTAGGTTGGGTCAAAATCAAACAAAAGAGTCTGGGC 2067

Db   31398 GATGAAAAGGCAAGCAAATTAGAGCCCTTTAGGTTGGTTCAAAATCAAACAAAAGAGTCTGGGC 31457

Qy    2068 CTGTGTCTTGTTGTGCCCTGGGGTGTCCTGGCTGACCTGCACTGGCAGGCACAGGATGAAG 2127

Db   31458 CTGTGTCTTGTTGTGCCCTGGGGTGTCCTGGCTGACCTGCACTGGCAGGCACAGGATGAAG 31517

Qy    2128 CGGAAGGAAAAAACTCACACTGACTTCCAAGACACCACAACCCCTCGTGTCTACATTACCTT 2187

Db   31518 CGGAAGGAAAAAACTCACACTGACTTCCAAGACACCACAACCCCTCGTGTCTACATTACCTT 31577

Qy    2188 GCCCAGGCCTCTCCAAGCTCAGCCATTGGGACCCCAGCTCAGCAGACCTCTGCTTCTGGGGCT 2247

Db   31578 GCCCAGGCCTCTCCAAGCTCAGCCATTGGGACCCCAGCTCAGCAGACCTCTGCTTCTGGGGC 31637

Qy    2248 GCTGACGGGCATCTGGCAAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGG 2307

Db   31638 GCTGACGGGCATCTGGCAAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGG 31697

Qy    2308 TTAGATGGGGGCCATCATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATC 2367

Db   31698 TTAGATGGGGGCCATCATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATC 31757

Qy    2368 CCTCTGAGCTGACAGTCTTCATTCACCTCTGTTTCCAAGAGTTCTTTTGCAGCAAATGTCCTAT 2427

Db   31758 CCTCTGAGCTACAGTCTTCATTCACCTCTGTTTCCAAGAGTTCTTTTGCAGCAAATGTCCTAT 31817

Qy    2428 GTCTTGGAGGATGAGAGGGGAGGTAAACATTCTAATTGCATCATAGATTTGGAAAAG 2487

Db   31818 GTCTTGGAGGATGAGAGGGGAGGTAAACATTCTAATTGCATCATAGATTTGGAAAAG 31877

Qy    2488 ACGCTAGAGAAGCATATGGAATACATGGCCCTGTTTGGGGCATCAACCACCACGTTTCCTATTG 2547

Db   31878 ACGCTAGAGAAGCATATGGAATACATGGCCCTGTTTGGGGCATCAACCACCACGTTTCCTATTG 31937
```

```
Qy    2548 GGCTGGAGGAAGTGATGAGGGGAGAGGAGATGGAGAAGCAATCTTTCACTCGCCGGCTGTGCT 2607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   31938 GGGCTGTTAAGTGATGAGGGGGGAGAGAGATGGAGAACATCTTTCACTCGCCGGCTGTCT 31997

Qy    2608 CAGGGGAGGAAACCTGATGCAGTGCAGTGGGGTCCCGTCCCTGCTGCTGTCAGCCACACTCT 2667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   31998 CAGGGGAGGAACCTGATGCAGTGGGGTCCCGTCCCTGCGAGCTGTCAGCCACACTCT 32057

Qy    2668 CTGGAGTCCCTCCACTGCTTGTTGTACGAGACTCGGAACAAAAACGTTCCTGACACAAGTGATG 2727
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32058 CTGGAGTCCCTCCACTGCTTGTTGTACGAGACTCGGAACAAAAACGTTCCTGACACAAGTGATG 32117

Qy    2728 GCCCCATTTCGAAGAAATGGGCATGTGTGTAGAAACAGAGAGTCTTAGTGTGCACT 2787
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32118 GCCCCATTTCGAAGAAATGGGCATGTGTGTAGAAACAGAGAGTCTTAGTGTGCACT 32177

Qy    2788 TTCTGCATTAAATTCAGCCGCCCCAGGTGAGAAGCTTCACTGATTGAGGGGCAGCCAGCAC 2847
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32178 TTCTGCATTAAATTCAGCCCCCCACCGTGAAGAAGCTTCAGCTGATTGAGGGGCAGGCAGCAC 32237

Qy    2848 AGATCAACAATGGAGCCCCCACCCAGTGGTAGTCCTGT 2881
           ||||||||||||||||||||||||||||||||||||||
Db   32238 AGATCAACAATGGAGCCCCCACCCATGGTAGTCCTGT 32271
```

```
RESULT 15
US-10-027-632-150437
; Sequence 150437, Application US/10027632
; Publication No. US20030240075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150437
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-150437
```

```
Db    121 CCTCGGATGGTGTAGAGATGAGCCAGGAGGATGGGGTTCTTGCAGGAGGAGGCCGAGCTCTGGTGCT 180
```

```
Qy   1806 GCACTGGAGCCAGCCACCAGCCCCGGACGGCCGGATGCAGTGCACTGCTGTGGGGCAGTGTGCTGGGAAAGCTAT 1865
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCACTGGAGCCAGCCACCAGCCCGGACGGCCGGATGCAGCAGCTGCTGGGGCAGTTTGCTGTGGGGAAAACTAT 240

Qy   1866 ACTTCCCGGACGGGCCATCCTTCCTGATCAGTCACGGCTCGGACCACAGCTCTTGCAGGAACCTCATTCC 1925
          ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db    241 ACTTCCCGGACGGGCCATCCTTCTGATCACGGGTCGGACCACAGCTCTTGCAGGAACCTCATTCC 300

Qy   1926 TTCTTTGGAGGCAGGGCACGTTGGGGTAGAGAGGTGCCTGGGGGTTCTCTGAGTCATCCAGCAGTGAAGGA 1985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTCTTTGGAGGCAGGGCACGTTGGGGTAGAGAGGTTCCTGGGGGTTCCAGCAGGAAGGA 360

Qy   1986 ATATTTCTACAGATATTTCACAGATGAAAGGCAACGACAAATTAGAGCCTTTAGGTTGGTTTCAA 2045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATATTTCTACAGATATTTCACAGATGAAAGGCAAGCAAATTAGAGCCTTTAGGTTGGTTCAA 420

Qy   2046 ATCAAAACAAAGAGCTCGGGCCCGTTGGGTGTCTTGTGTGGCCCCTGGGGTGTCCCTGGGCCTGGCAC 2105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATCAAACAAAGAGCTCTGGGCCCCGTTGGGTGTCTTGTGTGGCCCTGGGGTGTCCCTGGGCCTGGCAC 480

Qy   2106 TTGCCTGATCGCAGCAGATGAAAGCCGGAAGGAAGGAAAAAACTCACACTGACTTGAGTCCAAAGACCACCAC 2165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTGCCTGATCGCAGCAGATGAAAGCCGGAAGGAAGGAAAAAACTCACACTGACTTCAAAGACCACACCAC 540

Qy   2166 AACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAAGCTTCAGCCATTGGGACCCCCAGCT 2225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAAGCTTCAGCCATTGGGACCCCCAGCT 600

Qy   2226 CAGAGACCCTCTGCTCTCTCGGCGTTGCTGAGCGGCCATCTGGGCAAAAAAAGACCCTTTTCAGTCC 2285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGAGACCCTCTGCTCT-TGGCGTCGTGAGCGGCCATCTGGGCAAAAAAAGACCC--TTTCAGTCC 658

Qy   2286 AGATGACCTCAGGAAGCATGGGTTAGAATGGGGCCAGTCATCATCTCCACCTTCTTGAAGATGGG 2345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    659 AGATGACCTCAGGAAGCATGGGTTAGAATGGGGCCCTCATCTCCACCTTCTTGAAGAGGGG 718

Qy   2346 TA 2347
          ||
Db    719 TA 720
```

```

```
Qy   1626 CTGCAGAGAGCTGGCCCAGTCCAAGGTGTCGTGAGTCTCGCTCGAGTCTCATCGGGAAAAGATGG 1685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CTGCAGAGAGCTGGCCCAGTCCAAGGTGTGAGTCTCGCTCGAGTCTCATCGGGAAAAGATGG 60

Qy   1686 GACAGCCACTCCGGCGCTCCCATTAGACAGATCCTGTTCTAGGCCAGAGGCGGCTGCTGCTTCCTCAT 1745
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GACAGCCACTCCGGCGCTCCCATTAGACAGATCCTGTTCTAGGCCAGAGGCGGCTGCTTCTCAT 120

Qy   1746 CCTCGGATGGTGTAGAGATGAGCCAGGAGGATGGGGTTCTTGCAGGAGGAGGCCGAGTTCTGAGCTCTGTGCT 1805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
; Query Match            12.8%;  Score 694.4;  DB 13;  Length 720;
; Best Local Similarity  99.3%;  Pred. No. 4.1e-182;
; Matches  717; Conservative  1; Mismatches  2; Indels  2;  Gaps  2;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:06 ; Search time 1342.51 Seconds
                                    (without alignments)
                                    10946.442 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gcccaggggcctgggaggt.........aggaataagaagttacctac 5444

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

No.	Score	% Query Match	Length	DB ID	Description
1	5444	100.0	5523	22	AAF83651	Human CARD-7 polyp
2	5444	100.0	5523	25	ABS56030	cDNA encoding huma
3	5439.2	99.9	5523	25	ABS55497	cDNA encoding huma
4	4879.2	89.6	5100	24	AAL47127	Pyrin domain conta
5	4765.6	87.5	5122	22	AAD12951	Human G-protein co
6	4148	76.2	4422	22	AAD02760	Human NB-ARC and C
7	4078	74.9	4194	22	AAD02761	Human NB-ARC and C
8	3952	72.6	4329	22	AAD02762	Human NB-ARC and C

Result
No.   Score

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

					DB ID	Description

9   3400.4   62.5   4556   22   AAD02764   Human NAC beta iso
10 3210.4   59.0   4466   22   AAD02765   Human NAC gamma or
11 2572.6   47.3   2657   21   AAA78392   Human secreted pro
12   484.4   8.9   578   22   AAL01744   Human reproductive
13   484.4   8.9   578   23   ABL97037   Human testicular a
14   482.8   8.9   487   24   ABV97890   Human pancreatic c
15   414.4   7.6   416   24   ABV97891   Human pancreatic c
16   399.8   7.3   .3186   24   AAL44363   Human PYRIN-8 cDNA
17   398   7.3   3300   24   AAL47129   Pyrin domain conta
18   369.8   6.8   3857   24   AAD14323   Human PYRIN 1 (PYR
19   369.8   6.8   3857   22   ABX93556   Huma cDNA encoding
20   357.2   6.6   2763   25   ABQ75801   Human secreted pro
21   319.6   5.9   330   21   AAC08520   Human ORFX ORF1069
22   316.2   5.8   2691   24   AAC75514   Human ORFX ORF1069
23   313.4   5.8   2847   24   AAL44365   Human PYRIN-11 cDN
24   304.6   5.6   2847   25   AAS01487   Human secreted pro
25   304.6   5.6   2847   25   ABZ73494   Secreted protein-e
26   255.2   4.7   2199   24   AAL47130   Pyrin domain conta
27   254   4.7   2575   24   AAI70683   Human PYRIN-3 cDNA
28   254   4.7   2767   25   ABT16018   Human secreted pro
29   254   4.7   3172   24   ABN99366   Human secreted pol
30   252.4   4.7   3368   24   AAL44366   Human PYRIN-3 cDNA
31   246.8   4.5   2614   25   ABV72512   Nucleotide sequenc
32   241.8   4.4   1557   22   AAL47143   Pyrin domain conta
33   237   4.4   1557   24   AAK17979   Pyrin domain conta
34   237   4.4   237   24   ABA36666   NOVX related polyn
35   237   4.4   237   24   ABA36780   Nucleotide sequenc
36   237   4.4   237   22   AAK43859   Human nucleotide b
37   237   4.4   237   22   AAI49872   Probe #1858 used
38   237   4.4   237   23   AAS43511   Human liver single
39   237   4.4   237   22   ABS18088   Human genome-deriv
40   226.4   4.2   469   22   ABA58276   Human foetal liver
41   226.4   4.2   469   22   ABA27416   Probe #5582 for ge
42   226.4   4.2   469   22   AAK06371   Human brain expres
43   226.4   4.2   469   22   AAK32039   Human brain expres
44   226.4   4.2   469   22   AAI15917   Probe #5950 for ge
45   226.4   4.2   469   22   AAI37897   Probe #6583 used t

ALIGNMENTS

RESULT 1
AAF83651

ID   AAF83651 standard; cDNA; 5523 BP.
XX
AC   AAF83651;
XX
DT   23-JUL-2001   (first entry)
XX
DE   Human CARD-7 polypeptide encoding cDNA.
XX
KW   CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;
KW   autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;
KW   antibacterial; antiviral; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Location/Qualifiers
FT   CDS   523..4812
FT   /*tag= a
FT   /product= "CARD-7"
XX
PN   WO200130813-A1.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-US29796.
XX
PR   27-OCT-1999;   99US-0428252.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX

PI   Bertin J;
XX
DR   WPI; 2001-343340/36.
DR   P-PSDB; AAB62571.
XX
PT   Identifying a modulator of interaction between caspase recruitment
PT   domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
PT   comprises measuring the binding of CARD-7 and CARD-5 in the presence of
PT   the compound -
XX
PS   Disclosure; Fig 1A-D; 80pp; English.
XX
CC   The invention relates to identifying compounds, that modulate interaction
CC   between caspase recruitment domain (CARD)-7 and CARD-5. The method
CC   involves measuring the binding of CARD-7 and CARD-5 in the presence of
CC   the compound (an increase in the binding of CARD-7 to CARD-5 in the
CC   presence of the compound compared to the binding in the absence of the
CC   compound indicates that the compound is a modulator of CARD-7-CARD-5
CC   interaction). Modulators of CARD-7 and CARD-8 expression or activity can
CC   be used to treat or diagnose disorders such as cancers, bacterial or
CC   viral infections, autoimmune disorders (systemic lupus erythematosus,
CC   immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
CC   organ-specific autoimmunity including multiple sclerosis, Hashimoto's
CC   thyroiditis, or Grave's disease, Psoriasis, graft rejection, allergies.
CC   CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
CC   of cellular processes including cell growth and cell death. The present
CC   sequence represents a cDNA encoding the human CARD-7 polypeptide.
XX
SQ   Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;

Query Match   100.0%; Score 5444; DB 22; Length 5523;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

(Nucleotide sequence alignment, Query (Qy) against Database (Db), positions 1–1680, 100% identity.)

```
Qy  1681 GATGGGACAGCCACTCCGGCTCCCATTAGACACAGATCCTTCGTCTAGGCAGCGGCTGCTC 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 GATGGGACAGCCACTCCGGCTCCCATTAGACACAGATCCTTCGTCTAGGCAGCGGCTGCTC 1740

Qy  1741 TTCATCCTTCGATGGTGTAGATGAGGCCAGGATGGGTCTTGCAGAGAGCCGAGTTCTGAGCTC 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TTCATCCTTCGATGGTGTAGATGAGGCCAGGATGGGTCTTGCAGAGAGCCGAGTTCTGAGCTC 1800

Qy  1801 TGTCTGCACTGGAGCCAGCCACAGCCGCGGCGGGAAGCCAGTTGCTGCTGGGGGAAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 TGTCTGCACTGGAGCCAGCCACAGCCGCGGCGGGAAGCCAGTTGCTGCTGGGGGAAA 1860

Qy  1861 ACTATACTTCCCGAGGGCATCCTTCCTGATCACCGGCCCGGGATCAGCTGCAGAACCTC 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACTATACTTCCCGAGGGCATCCTTCCTGATCACCGGCCCGGGATCAGCTGCAGAACCTC 1920

Qy  1921 ATTCCTTCTTTGGAGCAGGCAGCGTTGGGGTAGAGGGTTCTTGAGTCCAGCAGG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ATTCCTTCTTTGGAGCAGGCAGCGTTGGGGTAGAGGGTTCTTGAGTCCAGCAGG 1980

Qy  1981 AAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCAATTAGAGCCCTTTAGGTTG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 AAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCAATTAGAGCCCTTTAGGTTG 2040

Qy  2041 GTCAAAATCAAAAAGAGACTTCTGGGGGTCCTGGGGGTTCTTGAGTCCAGCAGG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTCAAAATCAAAAAGAGACTTCTGGGGGTCCTGGGGGTTCTTGAGTCCAGCAGG 2100

Qy  2101 TGCACTTGGCCTGATGCAGCGATGAAGCCGGAAGAAAACTCACACTGACTTCCAAGACC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TGCACTTGGCCTGATGCAGCGATGAAGCCGGAAGAAAACTCACACTGACTTCCAAGACC 2160

Qy  2161 ACCACAACCCTGTCTACATTACCTTGCCCAGGCTCTCCAAGCTCAGCCATTGGGACCC 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ACCACAACCCTGTCTACATTACCTTGCCCAGGCTCTCCAAGCTCAGCCATTGGGACCC 2220

Qy  2221 CAGCTCAGAGACCTCTGCTCTCTGGGTGTCCTGAGGGGCATCTGGCCTCCAAGCTCATTT 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 CAGCTCAGAGACCTCTGCTCTCTGGGTGTCCTGAGGGGCATCTGGCCTCCAAGCTCATTT 2280

Qy  2281 AGTCCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCAATCTCCACCTTCTTGAAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 AGTCCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCAATCTCCACCTTCTTGAAG 2340

Qy  2341 ATGGGTATTCTTCAAGAGCACACCCCATCCCTCTGAGCTACAGCTTCACTCTTGTTTC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 ATGGGTATTCTTCAAGAGCACACCCCATCCCTCTGAGCTACAGCTTCACTCTTGTTTC 2400

Qy  2401 CAAGAGTTCTTTGCAGCAAATGTCCTATGTCTTGGAGGATGAGAAGGGGAGAGGGTAAAACAT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CAAGAGTTCTTTGCAGCAAATGTCCTATGTCTTGGAGGATGAGAAGGGGAGAGGGTAAAACAT 2460

Qy  2461 TCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATATGGAAATACATGGCCTGTTT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 TCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATATGGAAATACATGGCCTGTTT 2520

Qy  2521 GGGGCATCAACCACAGCGTTTCCTATTGGGCCTGTTAAGTGATGAGGGGAGAGAGATG 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 GGGGCATCAACCACAGCGTTTCCTATTGGGCCTGTTAAGTGATGAGGGGAGAGAGATG 2580

Qy  2581 GAGAACATCTTTCACTGCCGGCGTGTCTCAGGGGAGGAACCTGAGCAGTGGGGTCCCGGTCC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 GAGAACATCTTTCACTGCCGGCGTGTCTCAGGGGAGGAACCTGAGCAGTGGGGTCCCGGTCC 2640

Qy  2641 CTGCAGCTGCTGCGTGCAGGCACACTCTCTGGAGTCCCTCCACTGTTGTACGAGACTCCG 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 CTGCAGCTGCTGCGTGCAGGCACACTCTCTGGAGTCCCTCCACTGTTGTACGAGACTCCG 2700

Qy  2701 AACAAAACGTTTCCTGACACAAGTGATGGCCCATTTCGGAAGAAATGGGCAGTGTGTAGAA 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 AACAAAACGTTTCCTGACACAAGTGATGGCCCATTTCGGAAGAAATGGGCAGTGTGTAGAA 2760

Qy  2761 ACAGACACATGGGAGCTTCCTTAGTGTGCACTTTCGCATTAAATTCACCGCCGCCACGTGAAGAAG 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 ACAGACACATGGGAGCTTCCTTAGTGTGCACTTTCGCATTAAATTCACCGCCGCCACGTGAAGAAG 2820

Db  2821 CTTCAGCTGATTTGAGGGCAGGCAGCGCACACAGATCAACATGGCAGGTAGTCCTG 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2821 CTTCAGCTGATTTGAGGGCAGGCAGCGCACACAGATCAACATGGCAGGTAGTCCTG 2880

Db  2881 TTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCCGTCCTCAAGGTC 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2881 TTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCCGTCCTCAAGGTC 2940

Db  2941 ACCAGAAACCTGAAGGGACTGGACCTAAGTGGAAAACTCGGCTGACAGTCAGTGAAG 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2941 ACCAGAAACCTGAAGGGACTGGACCTAAGTGGAAAACTCGGCTGACAGTCAGTGAAG 3000

Db  3001 AGTCTTTTGTAAGACCCTGAGACGCCCTCCGCTGCCTCCTGGGAGACCCGTTGGCTGGC 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3001 AGTCTTTTGTAAGACCCTGAGACGCCCTCCGCTGCCTCCTGGGAGACCCGTTGGCTGGC 3060

Db  3061 TGTGGCCTCACAGCTGAGGACTGCAAAGGGACTCGCCTTTGGGGTTGAGGAGCCAACCAGACC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3061 TGTGGCCTCACAGCTGAGGACTGCAAAGGGACTCGCCTTTGGGGTTGAGGAGCCAACCAGACC 3120

Db  3121 CTGACCGAGCTGGACCTGAGCTTCAAATGTGCTCACGGATGCTGGAGCCAAAACACCTTTGC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3121 CTGACCGAGCTGGACCTGAGCTTCAAATGTGCTCACGGATGCTGGAGCCAAAACACCTTTGC 3180

Db  3181 CAGAGACTGAGACAGCCGAGACCTGCGAAGCTACAGCGACTGCGTCAGCTGCCTC 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3181 CAGAGACTGAGACAGCCGAGACCTGCGAAGCTACAGCGACTGCGTCAGCTGCCTC 3240

Db  3241 ACGTCTGACTGGTGCCCAGGACCTGGCCTTCTTGTCCTTCAGTGAACTTGAGTGATGAG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3241 ACGTCTGACTGGTGCCCAGGACCTGGCCTTCTTGTCCTTCAGTGAACTTGAGTGATGAG 3300

Db  3301 CTAGACCTGCAGCAGCAAACCTGGATGGACGTTGGCGTGGGACTGCTCTTGTGAGGGGCTC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3301 CTAGACCTGCAGCAGCAAACCTGGATGGACGTTGGCGTGGGACTGCTCTTGTGAGGGGCTC 3360

Db  3361 AGGGCATCCTGCCCTGCAAACTCATACGGCCTGGGGGTTGGCCAGACAACTCTAGTGATGATGAG 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3361 AGGGCATCCTGCCCTGCAAACTCATACGGCCTGGGGGTTGGCCAGACAACTCTAGTGATGATGAG 3420

Db  3421 ATGAGGCAGGAACTGAAGGGCCCTGGAGCAGGAGAGGAAAACCTTCCCAATTGCTGAGGAAGGAA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3421 ATGAGGCAGGAACTGAAGGGCCCTGGAGCAGGAGAGGAAAACCTTCCCAATTGCTGAGGAAGGAA 3480

Db  3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGGCTGGATACGGAGGAGGATGAGTAATAGC 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGGCTGGATACGGAGGAGGATGAGTAATAGC 3540

Db  3541 ACATCCCTCACTCAAGCGGCAGGAGGATCGGATTCAGAGGGCGGCGGCTTCCCATGTTGCTCAG 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3541 ACATCCCTCACTCAAGCGGCAGGAGGATCGGATTCAGAGGGCGGCGGCTTCCCATGTTGCTCAG 3600

Db  3601 GCTAATCTCAAAACTCCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3601 GCTAATCTCAAAACTCCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA 3660

Db  3661 AGCTCCCCAGAGGTAGTACCGGGTGGAAACTCTTGTGCGGCTTCTCGCTGCCTTCCTCAAGGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3661 AGCTCCCCAGAGGTAGTACCGGGTGGAAACTCTTGTGCGGCTTCTCGCTGCCTTCCTCAAGGG 3720

Db  3721 GACCTTGCATACGGAAGCCCTTTGGGGACTGACGATGACTTCTGAGGGCCCCACAGGGGGCTGTG 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3721 GACCTTGCATACGGAAGCCCTTTGGGGACTGACGATGACTTCTGAGGGCCCCACAGGGGGCTGTG 3780

Db  3781 GCTACTGAGGTAGTTGACAAAGAAAAGGAACTTGTACCGGAGTTCACTTCCGTACTGGC 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3781 GCTACTGAGGTAGTTGACAAAGAAAAGGAACTTGTACCGGAGTTCACTTCCGTACTGGC 3840

Db  3841 TCCTACCCGTGGGCCCAAACACGGGGTCTCGCTTTTGTTGATGAGGAAGAGCGGGTGACCGTTGAG 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3841 TCCTACCCGTGGGCCCAAACACGGGGTCTCGCTTTTGTTGATGAGGAAGAGCGGGTGACCGTTGAG 3900
```

```
Db   3841  TCCTACCGCTGGCCCAACACGGGGTCTCTGCTTTGTGTGATGAGGAAGAAGCGGGTTGACCGTTGAG
Qy   3901  ATTGAATTCTCTGTGTTGTGTGGGGAGGACCAGTTTCCTGGGGTGAGATCAACCACAGCACAGCTGGATG  3960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901  ATTGAATTCTCTGTGTTGTGTGGGGAGGACCAGTTTCCTGGGGTGAGATCAACCACAGCACAGCTGGATG  3960
Qy   3961  GTGGGCAGGGGCCTTCTGCTTGGGACATCAAGGACCTGGAGCTGTGTGGAAGCTTGTCACCTC  4020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961  GTGGGCAGGGGCCTTCTGCTTGGGACATCAAGGACCTGGAGCTGTGTGGAAGCTGTCACCTC  4020
Qy   4021  CCTCACTTTGTGGCTCTTCCAAGGGGGCCCATGTGGACACATCCCTGTTCCAAATGGCCCAC  4080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021  CCTCACTTTGTGGCTCTTCCAAGGGGGCCCATGTGGACACATCCCTGTTCCAAATGGCCCAC  4080
Qy   4081  TTTAAAGAGAGGAGGGGATGGCTCTTCCCAAGGGGTCCTGGGGAGAGCCAGGGGTGGAGCTGTCACACATAGTT  4140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081  TTTAAAGAGAGGAGGGGATGGCTCTTCCCAAGGGGTCCTGGGGAGAGCCAGGGGTGGAGCTGTCACACATAGTT  4140
Qy   4141  CTGGAAAACCCCAGCTTCTCTCCCCCTTGGGAGTCCTCTGGAGAAGCCAGGCCAGGGTGGAGCTGCATCACATAATGCCCTG  4200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141  CTGGAAAACCCCAGCTTCTCTCCCCCTTGGGAGTCCTCTGGAGAAATGATCCATAATGCCCTG  4200
Qy   4201  CGCTTCATTCCCGTCACCTCTGTTGGTGTTGCTTTACCACCGGGTCCATCCTGAGGAAGTC  4260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201  CGCTTCATTCCCGTCACCTCTGTTGGTGTTGCTTTACCACCGGGTCCATCCTGAGGAAGTC  4260
Qy   4261  ACCTTCACCTCTACCTGATCCCAAGTGACTGCTCCATTCGGAAGAAACTGGAGCTCTGC  4320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261  ACCTTCACCTCTACCTGATCCCAAGTGACTGCTCCATTCGGAAGAAACTGGAGCTCTGC  4320
Qy   4321  TATCGGAAGCCCTGGGAGAAGACCAAGCTGTTCTCGGGGTTCTACGTTGGCCACTTGGGATCA  4380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321  TATCGGAAGCCCTGGGAGAAGACCAAGCTGTTCTCGGGGTTCTAGGTTGGCCACTTGGGATCA  4380
Qy   4381  GGGATCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTGGTGTGTGGGAGGCCTTGGTG  4440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381  GGGATCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTGGTGTGTGGGAGGCCCTTGGTG  4440
Qy   4441  AAACCAGGAGAATCTCATTGCCTCGACTACTCTGATCCCCTCCAGCCCGCCGCATAGCCGTTACCT  4500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441  AAACCAGGAGAATCTCATTGCCTCGACTACTCTGATCCCCTCCAGCCCGCCGCATAGCCGTTACCT  4500
Qy   4501  TCACCTCTGGATGCCCCGGCAGTTGCTGCACTTTGTGGACTATCGAGAGCAGCTGATA  4560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501  TCACCTCTGGATGCCCCGGCAGTTGCTGCACTTTGTGGACTATCGAGAGCAGCTGATA  4560
Qy   4561  GCCCGGAGTGACATCGGTGGAGGTTGTCTTGGACACTGCATGGACAGGTGCTGAGCCAG  4620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561  GCCCGGAGTGACATCGGTGGAGGTTGTCTTGGACACTGCATGGACAGGTGCTGAGCCAG  4620
Qy   4621  GAGCAGTACGAGAGGGTGCTGAGAACACGAGGCCCCAGCCCAGGTGCGAAGCTGTTC  4680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621  GAGCAGTACGAGAGGGTGCTGAGAACACGAGGCCCCAGCCCAGGTGCGAAGCTGTTC  4680
Qy   4681  AGCTTGAGCCAGTCCTGGGACAGAGGGAAGTGGACTTACCAAGCCCTGAAGGAG  4740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681  AGCTTGAGCCAGTCCTGGGACAGAGGGAAGTGGACTTACCAAGCCCTGAAGGAG  4740
Qy   4741  ACCCATCCTCACCTCATTATGGAACTCTGGGAGAAGGGCAGCAAAAAGGGACTCCTCGCCA  4800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741  ACCCATCCTCACCTCATTATGGAACTCTGGGAGAAGGGCAGCAAAAAGGGACTCCTCGCCA  4800
Qy   4801  CTCAGCAGCTGAAGTATCAACACCAGCCCCTTCAGCCCTGCCTTTGGGCTGACCC  4860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801  CTCAGCAGCTGAAGTATCAACACCAGCCCCTTCAGCCCTGCCTTTGGGCTGACCC  4860
Qy   4861  TTCCTTTGGGTCTCCAGTTTCTTCTCTCTGCAAACAAGTTGCCATCTGGTTTTGCCTTCCAGCA  4920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861  TTCCTTTGGGTCTCCAGTTTCTTCTCTCTGCAAACAAGTTGCCATCTGGTTTTGCCTTCCAGCA  4920
Qy   4921  CTAAAGTAAATGGAACTTTGATGATGCCTTTGCTGTGGGGCATTATATGTGTCATGCCCAGGGGATG  4980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921  CTAAAGTAAATGGAACTTTGATGATGCCTTTGCTGTGGGGCATTATATGTGTCATGCCCAGGGGATG  4980
```

RESULT 2
ABS56030

```
Qy   4981  CCACAGGGGGGCCCCAGTCCAGGTGGCCTTAACACAGCATCTCCACGGGAATGTCCATCTTGGGAGCT  5040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981  CCACAGGGGGGCCCCAGTCCAGGTGGCCTTAACACAGCATCTCCACGGGAATGTCCATCTTGGGAGCT  5040
Qy   5041  GGCAAGAGACCCCTGCAGACCCTCATTAGAGCCTCATCTGTTGGCCATCAGCACAGCCAAGCCCTAGA  5100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041  GGCAAGAGACCCCTGCAGACCCTCATTAGAGCCTCATCTGTTGGCCATCAGCACAGCCAAGCCCTAGA  5100
Qy   5101  GCCCTCCGGATCCCATCCAGGCCCAAAGGAGGAATAGGAGGGACATGGAACCATTGCCTC  5160
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101  GCCCTCCGGATCCCATCCAGGCCCAAAGGAGGAATAGGAGGGACATGGAACCATTGCCTC  5160
Qy   5161  TGGCTGTGTCACGGGGTGGAGCCCCCAAAATTGGGGGTTCACGGTGGGAGGCGACCACGTGGGATTC  5220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161  TGGCTGTGTCACGGGGTGGAGCCCCCAAAATTGGGGGTTCACGGTGGGAGGCGACCACGTGGGATTC  5220
Qy   5221  TTTGGCTTTTGTACAGGAAGGATCTCACAAGAGCAAGCCAACAGAGACTAAAGGTGGGAAGAAGTTT  5280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221  TTTGGCTTTTGTACAGGAAGGATCTCACAAGAGCAAGCCAACAGAGACTAAAGGTGGGAAGAAGTTT  5280
Qy   5281  ATTCAGAAAATAAAGGAGTATCACAGCTCTTTAGAATTTGTCTAGCAGGCTTTCCAGTT  5340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281  ATTCAGAAAATAAAGGAGTATCACAGCTCTTTAGAATTTGTCTAGCAGGCTTTCCAGTT  5340
Qy   5341  TTTACCAGAAAACCCTTATTTAAATTTTTTTAACTTTAAATTTAAGAATTTAAAAAAAT  5400
            ||||||||||||||||||||||||||||||||||||||| ||| |||||||||||||||
Db   5341  TTTACCAGAAAACCCTATTATTTAAATTTTTTTAACTTTAAATTTAAGAATTTAAAAAAAT  5400
Qy   5401  ACAAAAAAGAAAAAATGAAAAAATGAAAAGGAAATAAGAAAGTTACCTAC  5444
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401  ACAAAAAAGAAAAAATGAAAAAATGAAAAGGAAATAAGAAAGTTACCTAC  5444
```

PR　15-AUG-2001; 2001US-0931071.
XX
PA　(BERT/) BERTIN J.
XX
PI　Bertin J;
XX
PI
DR　WPI; 2003-028967/02.
DR　P-PSDB; ABG71633.
XX
PT　Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT　CARD-7 and CARD-5 in presence of 'test compound, measuring their
PT　binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PT　altered -
XX
PS　Disclosure; Fig 1; 43pp; English.
XX
CC
CC　The present invention relates to methods of identifying compounds
CC　that regulate caspase activity using caspase recruitment domain-7
CC　(CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC　method for identifying a compound that modulates the interaction
CC　between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC　converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
CC　CARD-8 molecules are useful as modulating agents in regulating a
CC　variety of cellular processes including cell growth, cell death, and
CC　inflammation. The methods of the invention are useful for identifying
CC　compounds that have the ability to increase/decrease apoptosis, or
CC　comprise the ability to induce caspase activation. The methods are
CC　useful for treating a disorder associated with inappropriate apoptosis
CC　or inappropriate inflammation. The methods are useful for treating
CC　disorders associated with an undesirably low rate of apoptosis such
CC　as cancer (preferably follicular lymphoma, chronic myelogenous
CC　leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
CC　infections, autoimmune diseases caused by low levels of apoptosis
CC　(e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC　and arthritis). The methods are also useful for treating disorders with
CC　undesirably high rates of apoptosis such as human immunodeficiency
CC　virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC　amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC　muscular atrophy, various forms of cerebellar degeneration, anaemia
CC　associated with chronic disease, aplastic anaemia, chronic neutropaenia,
CC　myelodysplastic syndromes, myocardial infarction, stroke, and
CC　various inflammatory disorders (e.g. Crohn's disease, reactive
CC　arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC　psoriasis, graft rejection, allergic rhinitis, food allergies,
CC　conjunctivitis, glomerular nephritis, etc). The present sequence
CC　encodes human CARD-7.
XX
SQ　Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;

Query Match　　　　100.0%; Score 5444; DB 25; Length 5523;
Best Local Similarity　100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GCCCCAGGGCGCCTGAGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCCGGCTCCACTCT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCCCCAGGGCGCCTGAGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCCGGCTCCACTCT   60

Qy      61 GGGGTTCTGAAAGCCCCATTCCCTGCTCGCGGCTTCCTCCCCACCCTCTTCTCTCAGCCTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGGGTTCTGAAAGCCCCATTCCCTGCTCGCGGCTTCCTCCCCACCCTCTTCTCTCAGCCTT  120

Qy     121 GCAGGTCAAGGGTTGATCTCAGGAGTCCAGGACCCAGGGAGGAAGAAATCTGAGGAACA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCAGTCAAGGGTTGATCTCAGGAGTCCAGGACCCAGGGAGGAAGAAATCTGAGGAACA  180

Qy     181 CAGAACAGTGAGCGGCGTTGCCCACACACCCTCACCAGGAGGAGGAGGA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAGAACAGTGAGCGGCGTTGCCCACACACCCTCACCAGGAGGAGGAGGAGGAATCTGACCCTCAC  240

Qy     241 CCTCCCTGCCTGGCCCTGGACCCCATCCCAGGACCTTCCCTATCAGCTGACTTCTTCCAGT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTCCCTGCCTGGCCCTGGACCCCATCCCAGGACCTTCCCTATCAGCTGACTTCTTCCAGT  300

Qy     301 GTCTTGCAGGCCCCTTCTGGGCTCCCTCTGGGGCTCTCCCCTCCCCCTCCCCCTCCCCCTCTAT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GTCTTGCAGGCCCCTTCTGGGCTCCCTCTGGGGCTCTCCCCTCCCCCTCCCCCTCCCCCTCTAT  360

Qy     361 CGGGCGTCTATCTGTTAGGTGCCCTGGGATTTTATAAAACTGGGTTTCGGAAATGCTGAATAAGA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CGGGCGTCTATCTGTTAGGTGCCCTGGGATTTTATAAAACTGGGTTTCGGAAATGCTGAATAAGA  420

Qy     421 GACGGTAAGAGCCAAGGCAAAAGACAGCACTGTTTCTCTGCCTGCTGCCTGATACCCTCACCAC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GACGGTAAGAGCCAAGGCAAAAGACAGCACTGTTTCTCTGCCTGCTGCCTGATACCCTCACCAC  480

Qy     481 CTGGGGAACACATCCCCAGACACACCCTTCTTAACTCCGGGGGACACAGAGGTGGCGGGAGCCTGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGGGGAACACATCCCCAGACACACCCTTCTTAACTCCGGGGGACACAGAGGTGGCGGGAGCCTGG  540

Qy     541 GGCCGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGTTCAGCTT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGCCGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGTTCAGCTT  600

Qy     601 CTGCCGCCAATAAAGCGCACTCCAGGAGGTCTTCCGGTGGAGCGCAACACCCGCCTCAGCGCAGAG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTGCCGCCAATAAAGCGCACTCCAGGAGGTCTTCCGGTGGAGCGCAACACCCGCCTCAGCGCAGAG  660

Qy     661 AAGACCAGTGGCATGGAGGGTGGGCCTCGTACCTGGGTGGCTCAGTATGGGGAGGCGAGCGGGGCC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGACCAGTGGCATGGAGGGTGGGCCTCGTACCTGGGTGGCTCAGTATGGGGAGGCGAGCGGGGCC  720

Qy     721 TGGGACCTAGCCCTCCATACCTGGGAGCACAGAGGGGGCTGAGGTCACTGTGCGCCCAAGCC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGGACCTAGCCCTCCATACCTGGGAGCACAGAGGGGGCTGAGGTCACTGTGCGCCCAAGCC  780

Qy     781 CAGGAAGGGGCGAGGCCCACTCTCTCCCCCTCATTCCCCTACAGCCCCAAGTGAAACCCCACCCTGGGG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CAGGAAGGGGCGAGGCCCACTCTCTCCCCCTCATTCCCCTACAGCCCCAAGTGAAACCCCACCCTGGGG  840

Qy     841 TCTCCCCAGCCCAACCCCACCCCACCCGCAGTGCTAATCCCCTGGATCCATGAAATTGCCGGGCG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTCCCCAGCCCAACCCCACCCCACCCGCAGTGCTAATCCCCTGGATCCATGAAATTGCCGGGCG  900

Qy     901 GGGTGCACCCAGGGCTCAGAGAGGAAGGGTTTGAGACAGCTGCCTGACACATCTGACGC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGGTGCACCCAGGGCTCAGAGAGGAAGGGTTTGAGACAGCTGCCTGACACATCTGACGC  960

Qy     961 CGCTGGGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTTCCAAGGTTCCCCAGACCAT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CGCTGGGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTTCCAAGGTTCCCCAGACCAT  1020

Qy    1021 GAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCCACCACCATCCCACCAGCGGGGGGGGAGCTGG  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCCACCACCATCCCACCAGCGTGCTGGGGGAGCTGG  1080

Qy    1081 GGATCCCACCTCAGCCCTAGCACCCAGAGCAGAGGAGGGCTCCTGGGACCCAATGG  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GGATCCCACCTCAGCCCTAGCACCCAGAGCAGAGGAGGGCTCCTGGGACCCAATGG  1140

Qy    1141 CCTCTCGGATGAAAACGTTCAGAGAAATTTACTACACAGAAAATCAGAGAGAAGAGAGAGAAA  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCTCTCGGATGAAAACGTTCAGAGAAATTTACTACACAGAAAATCAGAGAGAAGAGAGAGAAA  1200

Qy    1201 TCAGAGAAAAGGCAGGCCCCCCCATGGGCGCGCGGTGGTTAGGAAGCGGCCCCCACGCGCACC  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCAGAGAAAAGGCAGGCCCCCCCATGGGCGCGCGGTGGTTAGGAAGCGGCCCCCACGCGCACC  1260

Qy    1261 AGCCTACAGCCCCACCCACCCCATGGGGAGCCTTCGTGAGAGGAAGCGGCCTCGTTCCACA  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AGCCTACAGCCCCACCCACCCCATGGGGAGCCTTCGTGAGAGGAAGCGGCCTCGTTCCACA  1320

Qy    1321 TGGCCCCTGGGAAAAATGAGGATTTTAACCAAAAAAATTCACACAGAGCTACTTCTTACAAAGA  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TGGCCCCTGGGAAAAATGAGGATTTTAACCAAAAAAATTCACACAGAGCTACTTCTTACAAAGA  1380

Qy    1381 CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAGAGCTGGCCTGATTAATGTGGAGGAG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAGAGCTGGCCTGATTAATGTGGAGGAG  1440
```

This page contains a two-column DNA sequence alignment between a query sequence (Qy) and a database sequence (Db), shown as paired rows with match bars (|||) between them and residue position numbers flanking each row.

Left column (positions ~1381–2520):

```
Db 1381 CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAGAGCTGGCCTGATTATGTGGAGGAG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1441 ...

Db 1441 AATCGAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGAAGA 1500
Qy 1441 ...
```

(The alignment continues in paired Qy/Db blocks through position 2520 in the left column and position 3600 in the right column.)

Right column (positions ~2461–3600):

```
Db 2461 TCTAAATTGCATCATAGATTTGGAAAAAGACGCTAGAAGGCATATAGGAATACATGGCCTGTTT 2520
...
Db 3541 ACATCCCTCACTCAAAGCGGGCAGAGACTCGGATCAGAGAGAGGGCGGCCTTCCCATGTTGCTCAG 3600
```

```
RESULT 3
ABS55497
ID   ABS55497 standard; cDNA; 5523 BP.
XX
AC   ABS55497;
XX
DT   09-JAN-2003  (first entry)
XX
DE   cDNA encoding human caspase recruitment domain-7 (CARD-7).
XX
KW   Human; caspase activity; caspase recruitment domain-7; CARD-7;
KW   caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta,
KW   pseudoICE; ICEBERG; cell growth; cell death; inflammation;
KW   apoptosis; caspase activation; cancer; follicular lymphoma;
KW   leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
KW   autoimmune disease; systemic lupus erythematosus; reactive arthritis;
KW   human immunodeficiency virus infection; HIV infection; ALS;
KW   Alzheimer's disease; Parkinson's disease; amylotrophic lateral sclerosis;
KW   myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
KW   insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
```

```
KW    graft rejection; allergic rhinitis; food allergy; conjunctivitis;
KW    glomerular nephritis; cytostatic; virucide; immunosuppressive;
KW    dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.
XX
OS    Homo sapiens.
XX
FH    Key             Location/Qualifiers
FT    CDS             523..4812
FT                    /*tag= a
FT                    /product= "CARD-7"
XX
PN    US2002128219-A1.
XX
PD    12-SEP-2002.
XX
PF    15-AUG-2001; 2001US-0931071.
XX
PR    27-OCT-1999; 99US-0428252.
XX
PA    (BERT/) BERTIN J.
PA    (ALNE/) ALNEMRI E S.
XX
PI    Bertin J, Alnemri ES;
XX
XX
DR    WPI; 2003-028968/02.
DR    P-PSDB; ABG71631.
XX
PT    Assays for identifying compound that modulates the interaction of
PT    caspase recruitment domain-8 with a CARD-8 ligand or a compound that
PT    modulates activity of CARD-8 -
XX
PS    Disclosure; Fig 1; 49pp; English.
XX
CC    The present invention relates to methods of identifying compounds
CC    that regulate caspase activity using caspase recruitment domain-7
CC    (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC    method for identifying a compound that modulates the interaction
CC    between CARD-8 and caspase-1; pseudo-interleukin-1 (IL-1) beta
CC    converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
CC    CARD-8 molecules are useful as modulating agents in regulating a
CC    variety of cellular processes including cell growth, cell death, and
CC    inflammation. The methods of the invention are useful for identifying
CC    compounds that have the ability to increase/decrease apoptosis, or
CC    comprise the ability to induce caspase activation. The methods are
CC    useful for treating a disorder associated with inappropriate apoptosis
CC    or inappropriate inflammation. The methods are useful for treating
CC    disorders associated with an undesirably low rate of apoptosis such
CC    as cancer (preferably follicular lymphoma, chronic myelogenous
CC    leukaemia, melanoma, colon cancer, lung carcinoma, etc) viral
CC    infections, autoimmune diseases caused by low levels of apoptosis
CC    (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC    and arthritis). The methods are also useful for treating disorders with
CC    undesirably high rates of apoptosis such as human immunodeficiency
CC    virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC    amylotrophic lateral sclerosis (AlS) retinitis pigmentosa, spinal
CC    muscular atrophy, various forms of cerebellar degeneration, anaemia
CC    associated with chronic disease, aplastic anaemia, chronic neutropaenia,
CC    myelodysplastic syndromes, myocardial infarction, stroke, and
CC    various inflammatory disorders (e.g. Crohn's disease, reactive
CC    arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC    psoriasis, graft rejection, allergic rhinitis, food allergies,
CC    conjunctivitis, glomerular nephritis, etc). The present sequence
CC    encodes human CARD-7.
XX
SQ    Sequence 5523 BP; 1350 A; 1529 C; 1481 G; 1163 T; 0 other;
```

Query Match　　　99.9%;　Score 5439.2;　DB 25;　Length 5523;
Best Local Similarity　99.9%;　Pred. No. 0;
Matches 5441;　Conservative　0;　Mismatches　3;　Indels　0;　Gaps　0;

```
Qy           1 GCCCCAGGGCCTGGAGGAGGAGGTCTGAAGAAAACCTGGGAGACGCCAGCACGCCCGGGGGGGGGGGCTTCCACTCT   60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCCCCAGGGCCTGGAGGAGGAGGTCTGAAGAAAACCTGGGAGACGCCAGCACGCCCGGGGGGGGGGGCTTGACTCT   60

Qy          61 GGGGTTCTGAAAGCCCATTCCCTGCTTCTGCGGGCTCCTCCCACCCCACCTCTTCTTCTCAGCCTT  120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGGGTTCTGAAAGCCCATTCCCTGCTTCTGCGGGCTCCTCCCACCCCACCTCTTCTTCTCAGCCTT  120

Qy         121 GCAGCTCAAGGGGTTGGATCTCAGGAGTCCAGGACTCCCTCCTCACCCACCCCACCTCTTGAGGAACA  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCAGCTCAAGGGGTTGGATCTCAGGAGTCCAGGACTCCCTCCTCACCCACCCCACCTCTTGAGGAACA  180

Qy         181 CAGAAACAGTGAGCGTTGCCCACCACACCCCATCTCCCCTCACCACCATCTCCCCTCACCCTCAC  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTCTTGCAGGACGCCCCTCCTCGGGCTCCTCCCCCTGGCTTTTCCTACCACTCCCCCTCTAT  360

Qy         241 CCTCCCTGCCTGGGCCTGGACCCCATCCAGGAGACTCCCTATCAGCTGACTTCTTCCAGT  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCTCCCTGCCTGGGCCTGGACCCCATCCAGGAGACTCCCTATCAGCTGACTTCTTCCAGT  300

Qy         301 GTCTTGCAGGACGCCCCTCCTCGGGCTCCTCCCCCTGGCTTTTCCTACCACTCCCCCTCTAT  360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTCTTGCAGGACGCCCCTCCTCGGGCTCCTCCCCCTGGCTTTTCCTACCACTCCCCCTCTAT  360

Qy         361 CGGGCCTCTATCTGTTAGGGTGCCCTGGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA  420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGGGCCTCTATCTGTTAGGGTGCCCTGGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA  420

Qy         421 GACGGTAAGAGCCAAAGGCAAAGGACAGCACTGTTCTTCTGCCTGATACCCTCACCAC  480
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GACGGTAAGAGCCAAAGGCAAAGGACAGCACTGTTCTTCTGCCTGATACCCTCACCAC  480

Qy         481 CTGGGAACATCCCCCAGACACCCCTCTTAACTCCCGGACAGAGATGGCTGGCGGGAGCCTGG  540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGGGAACATCCCCCAGACACCCCTCTTAACTCCCGGACAGAGATGGCTGGCGGGAGCCTGG  540

Qy         541 GGCGCGGCCTCGGTTACTTGGAGTTCCTGGAGTTCCTGGAGGAGGAGGAGGAGGAGTTCAGCTT  600
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GGCGCGGCCTCGGTTACTTGGAGTTCCTGAGAGGAGGAGGAGGAGGAGTTCAGCTT  600

Qy         601 CTGCTTCGCCAATAAAGCGCACTCCAGGAGCTCTTGGGGTGAGACACCCGCTCAGCCAGAG  660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CTGCTTCGCCAATAAAGCGCACTCCAGGAGCTCTTGGGGTGAGACACCCGCTCAGCCAGAG  660

Qy         661 AAGACGAGTGGCATGGAAGTGGCCTTCGTGGCCTCGTTGGCCTCAGTATGGGGAGGCACGGGCC  720
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGACGAGTGGCATGGAAGTGGCCTTCGTGGCCTCGTTGGCCTCAGTATGGGGAGCAGGGGCC  720

Qy         721 TGGGACCTTAGCCCTTCCATACCTGGGGAGCAGATGGGGGTCACTGTGTCGCGCCCAAGCC  780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TGGGACCTTAGCCCTTCCATACCTGGGAGCAGATGGGGGTCACTGTGTCGCGCCCAAGCC  780

Qy         781 CAGGAAGGGGCAGGCCACTCTCCCCTACAGCCCAAGTAGTGAAACCCCACCTGGGG  840
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CAGGAAGGGGCAGGCCACTCTCCCTCATTCCCCTACAGCCCAAGTAGTGAAACCCCACCTGGGG  840

Qy         841 TCTCCCAGCCAACCCACCTCCACCCCGCAGTGCTAATGCCCTGGGAGCCATGAATTGCCGGCG  900
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TCTCCCAGCCAACCCACCTCCACCCCGCAGTGCTAATGCCCTGGGATCCATGAATTGCCGGCG  900

Qy         901 GGGGTGCACCCAGGGGCTCAGAGGAGGAAGGGTTTTGAGGACAGCTGCCTGACACAATCTGGACGC  960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GGGGTGCACCCAGGGGCTCAGAGGAGGGGTTTTGAGGACAGCTGCCTGACACAATCTGGACGC  960

Qy         961 CGCTGGAGAGAAAATCTCTGGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCCAGACCAT  1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CGCTGGAGAGAAAATCTCTGGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCCAGACCAT  1020

Qy        1021 GAGTCTCCAAGCCACAGGAGTCAACGCCCCCACATCCAACCAGTGCTTGTGGGGAGCTGG  1080
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GAGTCTCCAAGCCACAGGAGTCAACGCCCCCACATCCAACCAGTGCTTGTGGGGAGCTGG  1080

Qy        1081 GGATCCCCACCCTCAGCCCCAGCCTAGCCACCCAGAGAGCAGCAGGAGGCTTCCTGGACCCAATGG  1140
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GGATCCCCACCCTCAGCCCCAGCCTAGCCACCCAGAGAGCAGCAGGAGGCTTCCTGGACCCAATGG  1140
```

Qy   1141 CCTTCTGGATGAAAACGTCAGGGAATTACTACACAGAGAAATCAGAGAGAGAGAGAGAGAA 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCTTCTGGATGAAAACGTCAGGGAATTTACTACACAGAAATCAGAGAAAGAGAGAGAGAAA 1200

Qy   1201 TCAGAGAAAAGGCAGGCCCCCCATGGGCAGCGGTGGTAGGAACGCCCCCACACGGCCACACC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGGAGAAAAGGCAGGCCCCCATGGGCAGCGGTTAGGAACGCCCCACAGGGCACACC 1260

Qy   1261 AGCCTACAGCCCCCACCACCCCATGGGAGGCCTTCGTGAGAGAGCCTCTGTTCCACA 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGCCTACAGCCCCCACCACCCCATGGGAGGCCTTCGTGAGAGAGCCTCTGTTCCACA 1320

Qy   1321 TGGGCCTTGGAAAAATGAGGGATTTTAACCAAAAAATTCACACAGTGCTACTTCTACAAAGA 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGGGCCTTGGAAAAATGAGGGATTTTAACCAAAAAATTCACACAGTGCTACTTCTACAAAGA 1380

Qy   1381 CCTTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGAAGCCTGATTATTTACACAGGAGG 1440

Db   1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGAAA 1500

[DNA sequence alignment — Qy (query) and Db (database) paired rows continue through the page, positions 1141–3360]

```
Qy   3301  CTAGACCTGCAGCAGAACAAACCTTGGATGACGTTGGCGTGGCGACTGCTCTGTGAGGGGCTC  3360
Db   3301  AGGCATCCTGCTTGCAAATCTATACGCCTCTGACCAGAACAACTTCTGAGTGATGAG       3360
Qy   3361  AGGCATCCTGCTTGCAAATCTATACGCCTCTGACCAGAACAACTTCTGAGTGATGAG  3420
Db   3361  AGGCATCCTGCCTGCAAACTCATACGCCTGGGGCTGGGACCAGCAACTCGAGTGGATGAG  3420
Qy   3421  ATGAGGGCAGGAGACTGAGGGCCCTGGAGCAGGACAGAGAAAACCTCAGCTGCTCATCTTCAGCAGA  3480
Db   3421  ATGAGGGCAGGAGACTGAGGGCCCTGGAGCAGGAGACAAACCTCAGCTGCTCATCTTCAGCAGA  3480
Qy   3481  CGGAAAACCAAGTGTGATGATGACCCCTACTGAGGGGGCCTGGATACGGGGAGAGATGAATAGC  3540
Db   3481  CGGAAAACCAAGTGTGATGATGACCCCTACTGAGGGGGCTGGATACGGGGAGAGATGAATAGC  3540
Qy   3541  ACATCCTCACTCAAGCGCGCAGAGACTCGGATCAGAGGAGGGCGGCGCCTTCCCATGTTGCTCAG  3600
Db   3541  ACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGGAGGGGCCTTCCCATGTTGCTCAG  3600
Qy   3601  GCTAATCTCAAAACTTCCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA  3660
Db   3601  GCTAATCTCAAAACTTCCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA  3660
Qy   3661  AGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGGTGCGTTGGTGCCTTCCTCGCCTCTCAAGGG  3720
Db   3661  AGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCCTGCCCTCTCAAGGG  3720
Qy   3721  GACCTGCATACGAAGCCTTTGGGGACTGACGATGACTCTTGGGGGGCCCTGTG  3780
Db   3721  GACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGGGGGGCCCTGTG  3780
Qy   3781  GCTACTGAGGTAGTTGACAAAGAAAAAGAACTTGTACCGGAGTTCACTTCCCGTAGCTGGC  3840
Db   3781  GCTACTGAGGTAGTTGACAAAGAAAAGAACTGTACCGAGTTCACTTCCCGTAGCTGGC  3840
Qy   3841  TCCTACCGCTGGCCCACACGGGTCTCTGCTTTGTGATGAGAGAAGCGGTGACCGTTGAG  3900
Db   3841  ATTGAATTCTGTGTGTGGGACCAGTTCCTTGGGGACCACACCACTGGATG  3900
Qy   3901  ATTGAATTCTGTGTGTGGGACCAGTTCCTTGGGGACCACACCACTGGATG  3960
Db   3901  GTGGCAGGGCCCTCTGCTGGACATCAAGGCTGACCTGAGCTGTGCCACCTC  3960
Qy   3961  GTGGCAGGGCCCTCTGCTGGACATCAAGGCTGACCTGAGCTGTGCCACCTC  4020
Db   3961  CCTCACTTTGTGGCTCTCCAAGGGGGCCCATGTGGACACATCCCTGTTCCAAATGGCCCAC  4080
Qy   4021  CCTCACTTTGTGGCTCTCCAAGGGGGCCCATGTGGACACATCCCTGTTCCAAATGGCCCAC  4080
Db   4021  TTTAAAGAGGAGGGGGATCTCCTGGAGGAAGCCAGCCAGGCTGTCATCACTAGTT  4140
Qy   4081  TTTAAAGAGGAGGGGGATGCTCCTGGAGAAGCCAGCCAGGCTGTCATCACTAGTT  4140
Db   4081  CTGGAAAAACCCCAGCTTCTCCCCCCTTTGGGAGTCCTCCTGAAAAATGATCCATAATGCCCTG  4200
Qy   4141  CTGGAAAAACCCCAGCTTCTCCCCCCTTTGGGAGTCCTCTGAAAATGATCCATAATGCCCTG  4200
Db   4141  CGCTTCATCCCGTGTGGTTGCTTACCACCGGTCCATCCTGAGGGAGTC  4260
Qy   4201  CGGCTTCATCCCGTGTGGTTGCTTACCACCGGTCCATCCTGAGGGAGTC  4260
Db   4201  ACCTTCCACCTCTACCTGATCCAAGTGACTGACTGTCCAGTGGAGCTCTGC  4320
Qy   4261  ACCTTCCACCTCTACCTGATCCAAGTGACTGACTGTCCAGTGGAGCTCTGC  4320
Db   4261  TATCGGAAGCCCTGAGAGAGACCAGCTGTTCTCGGAGTTCTTACGTTGGCTCCACTTGGGATCA  4380
Qy   4321  TATCGGAAGCCCTGAGAGAGACCAGCTGTTCTCGGAGTTCTTACGTTGGCTCCACTTGGGATCA  4380
Db   4321  GGGATCAGGCTGCAAGTGANAAGACAAGAGAATCGAGACTCTGGTGTGTGGGGAGGGCTTGGTG  4440
Qy   4381  GGGATCAGGCTGCAAGTGANAAGACAAGAGAATCGAGACTCTGGTGTGTGGGGAGGGCTTGGTG  4440
Db   4381  GGGATCAGGCTGCAAGTGANAAGACAAGAGAATCGAGACTCTGGTGTGTGGGGAGGGCTTGGTG  4440

Qy   4441  AAAACCAGGAGGATCTCATGCCTGCAACTACTCTGATCCCTCCCAGCCCAGCCGGTACCT  4500
Db   4441  AAAACCAGGAGGATCTCATGCCTGCAACTACTCTGATCCCTCCCAGCCCAGCCGGTACCT  4500
Qy   4501  TCACCTCTGGATGCCCCCGCAGTTGTGCTGCACTTTGTGGAGACCAGTATCGAGAGCAGCTGATA  4560
Db   4501  TCACCTCTGGATGCCCCCGCAGTTGTGCTGCACTTTGTGGACCAGTATCGAGAGCAGCTGATA  4560
Qy   4561  GCCCGGAGTGACATGGGTGGAGGGTGGAGGTTGTCTTGGACAAACTGCATGGACAGGTGCTGAGCCAG  4620
Db   4561  GCCCGGAGTGACATGGGTGGAGGGTGGAGGTTGTCTTGGACAAACTGCATGGACAGGTGCTGAGCCAG  4620
Qy   4621  GAGCAGTAGGAGAGGGTGCTGCTGAGAACAGAGGCCCAGCAGGCGGAAGCTGTTC  4680
Db   4621  GAGCAGTAGGAGAGGGTGCTGGCTGAGAACAGAGGCCCAGCAGGCGGAAGCTGTTC  4680
Qy   4681  AGCTTGACCCAGTCTGGGACCCGGAAGGTGCAAAGAGCTCTCAGCACCCCTGAAAGGAG  4740
Db   4681  AGCTTGAGCCAGTCCTGGGAACCGGAAGTGCAAAGATGGGACTCTACCAAGGCCCTGAAAGGAG  4740
Qy   4741  ACCCATCCTCACCTCATTTATGGAACTCTGGGAGGAGAAGGGCAGCAAAAAGGGAGACTCCTGCCA  4800
Db   4741  ACCCATCCTCACCTCATTTATGGAACTCTGGGAGGAGAAGGGCAGCAAAAAGGGAGACTCCTGCCA  4800
Qy   4801  CTCAGCAGCTGAAGTATCAAACACCAGCCCCTAACAGCATCTCAGGGCATTATGTGTCCATCAGGGGATG  4980
Db   4801  CTCAGCAGCTGAAGTATCAAACACCAGCCCCTAACAGCATCTCAGGGCATTATGTGTCCATCAGGGGATG  4980
Qy   4921  CTAAAGTAAATGGAACTTTGAATGAATGCCCTTTGCTCGGGCATTATGTGTCCATCAGGGGATG  4980
Db   4921  CTAAAGTAAATGGAACTTTGAATGAATGCCCTTTGCTCGGGCATTATGTGTCCATCAGGGGATG  4980
Qy   4981  CCACAGGGGGGCGCCCCAGTCCAGTGGCCTAACAGCACTCTCAGGGAATGTCCATCTGGAGCT  5040
Db   4981  CCACAGGGGGGCGCCCCAGTCCAGTGGCCTAACAGCACTCTCAGGGAATGTCCATCTGGAGCT  5040
Qy   5041  GGCAAGACCCCTGCAGACCCTCATAGAGCCCAAGAGCCAAGCCTAGA  5100
Db   5041  GGCAAGACCCCTGCAGACCCTCATAGAGCCCAAGAGCCAAGCCTAGA  5100
Qy   5101  GCCCTCCGGATCCCATCCAGGCGCCAAAGAGGAATAGGAGGGACATGGAACCCATTTGCCTC  5160
Db   5101  GCCCTCCGGATCCCATCCAGGCGCCAAAGAGGAATAGGAGGGACATGGAACCCATTTGCCTC  5160
Qy   5161  TGGCGTGTGTCACAGGGTGAGCCCCCAAATTGGGGGTTCAGCGTGGGAGGCCACGTGGATTC  5220
Db   5161  TGGCGTGTGTCACAGGGTGAGCCCCCAAATTGGGGGTTCAGCGTGGGAGGCCACGTGGATTC  5220
Qy   5221  TTGGCTTTTGTACAGGAAGATCTACAGAGCAAGCCAACAGAGTAAAGTGGAAGAAGTTT  5280
Db   5221  TTGGCTTTTGTACAGGAAGATCTACAGAGCAAGCCAACAGAGTAAAGTGGAAGAAGTTT  5280
Qy   5281  ATTCAGAAAAATAAAGGAGTATCACAGCTCTCTTTTGAAATTTGTTCTTAGCGCAGGCTTTTTCCAGTT  5340
Db   5281  ATTCAGAAAAATAAAGGAGTATCACAGCTCTCTTTTGAAATTTGTTCTTAGCGCAGGCTTTTTCCAGTT  5340
Qy   5341  TTTACCAGAAAAATAAAGGAATAAGGAATTAAGGAAGTTACCTAC  5400
Db   5341  TTTACCAGAAAAATAAAGGAATAAGGAATTAAGGAAGTTACCTAC  5400
Qy   5401  ACAAAAAAGAAAAAAATGAAAAATAAAGGAATTAAGGAAGTTACCTAC  5444
Db   5401  ACAAAAAAGAAAAAAATGAAAAATAAAGGAATTAAGGAAGTTACCTAC  5444
```

RESULT 4
AAL47127
ID   AAL47127   standard; DNA; 5100 BP.

```
XX
AC   AAL47127;
XX
DT   20-AUG-2002 (first entry)
XX
DE   Pyrin domain containing protein NALP1-hs coding sequence.
XX
PF   Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW   antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW   neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW   nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW   inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW   arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW   osteoarthritis; glomerulonephritis; gene; ds.
XX
OS   Unidentified.
XX
PN   WO200240668-A2.
XX
PD   23-MAY-2002.
XX
PF   30-OCT-2001; 2001WO-EP12545.
XX
PR   15-NOV-2000; 2000DE-1056687.
PR   30-NOV-2000; 2000DE-1059595.
XX
PT   New DNA encoding protein with pyrin domain, useful for treating
PT   diseases involving impaired signal transduction, particularly
PT   inflammation, also proteins and antibodies -
XX
PA   (APOT-) APOTECH RES & DEV LTD.
XX
PI   Tschopp J, Martinon F;
XX
DR   WPI; 2002-427093/45.
DR   P-PSDB; AAO17855.
XX
PS   Claim 5; Fig 1; 116pp; German.
XX
CC   The present invention relates the DNA and their encoded proteins, where
CC   the proteins contain at least one PYD (pyrin) domain. These can be used
CC   to treat diseases associated with impaired intracellular signal
CC   transduction, particularly inflammation such as psoriasis,
CC   arteriosclerosis, bacterial or viral infections (particularly meningitis
CC   and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC   sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC   and Parkinson's diseases. The present sequence is a coding sequence of
CC   the invention.
XX
SQ   Sequence 5100 BP; 1251 A; 1376 C; 1404 G; 1069 T; 0 other;

Query Match             89.6%;  Score 4879.2;  DB 24;  Length 5100;
Best Local Similarity   99.4%;  Pred. No. 0;
Matches 4896;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy    516 GACAGAAGATGGCTGCGGAGACCTGGGAGCCGCCGGCCTGGGCCTGTTGTACTTGGAGTTCCTGAAGAA 575
          ||||| |||||||| ||| |||||| |                    ||||||||||||||||||||||
Db    153 GAACAAGATGGCTCGGCGAGACGAGTTCAGCGTCTTCGCCGGCGCCCTTGGGCCTGTTACTTGGAGTTCCTGAAGAA 212

Qy    576 GGAGGAGCTGAAGGAGTTCAGCGTCTTCGCCGGCGCCCTTGGGCCTGTTACTTCCAGGAGCTCTTC 635
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 GGAGGAGCTGAAGGAGTTCAGCGTCTTCGCCGGCGCCCTTGGGCCTGTTACTTCCAGGAGCTCTTC 272

Qy    636 GGGGTGAGAGACACCCCGCTCAGCCAGAGAAGACGAGAGTGGCATGGAGGCCTCGTACCTCTGGT 695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 GGGGTGAGAGACACCCCGCTCAGCCAGAGAAGACGAGAGTGGCATGGAGGCCTCGTACCTCTGGT 332

Qy    696 GGCTCAGTATGGGAGGACAGCCAGCGCGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGAGGG 755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 GGCTCAGTATGGGAGGACAGCCAGCGCGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGAGTGC 392

Qy    756 GCTGAGGTCACTGTGCGCGCCCAAGCCCAGGAAGGGGCAGGCCCACTCTCCCGTCCATTCCCCTA 815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 GCTGAGGTCACTGTGCGCGCCCAAGCCCAGGAAGGGGCAGGCCCACTCTCCCGTCCATTCCCCTA 452

Qy    816 CAGCCCAAGTGAAACCCCACCTTGGGGTTCTCCCAGCCACCCTCCACCGCGCAGTGCTAAT 875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 CAGCCCAAGTGAAACCCCACCTTGGGGTTCTCCCAGCCACCCTCCACCGCGCAGTGCTAAT 512

Qy    876 GCCCTGGATCCATGAATTGCCGGCGGGGTGCACCCAGCCAACCCTCCACTCCTCTACCA 935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    513 GCCCTGGATCCATGAATTGCCGGCGGGGTGCACCCAGCCTCAGGAGAGGGTTTTTGAG 572

Qy    936 ACAGCTGCCTGACACATCTGGACGCCGCTGGGAGGAGAGAAATCTCTGCCTCACTCCTCTACCA 995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    573 ACAGCTGCCTGACACATCTGGACGCCGCTGGGAGGAGAGAAATCTCTGCCTCACTCCTCTACCA 632

Qy    996 AGCTCTTCCAAGCTCCCCAGACCAATGAGTCTTCCAAGCCACCCAGAGTCACCCAAACGCCCCCAC 1055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    633 AGCTCTTCCAAGCTCCCCAGACCAATGAGTCTTCCAAGCCACCCAGAGTCACCCAAACGCCCCCAC 692

Qy   1056 ATCCACAGCAGTGTCTGTGGGAGGACTGGGGACATCCCACCTCAGCCCAGCCTAGCACCCAGAGA 1115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    693 ATCCACAGCAGTGTCTGTGGGAGGACTGGGGATCCCACCTCAGCCCAGCCTAGCACCCAGAGA 752

Qy   1116 GCAGGAGAGGGCTTCCTGGGGACCCCAATGGGCCTTCTGGATGAAACGTTCAGGAAATTTACTACACAGA 1175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    753 GCAGGAGAGGGCTTCCTGGGGACCCCAATGGGCCTTCTGGATGAAACGTTCAGGAAATTTACTACACAGA 812

Qy   1176 AATCAGAGAAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGAGACCCAGAGTGCCCCATGGGGCAGCCAGCCGGTGGT 1235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    813 AATCAGAGAAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGAGACCCAGAGTGCCCCATGGGGCAGCCAGCCGGTGGT 872

Qy   1236 AGGAAACGCCCCCACAGGGCGCACACCGGCCTACAGCCCCACGCCCCACCCACCCCATCCCAGCCCTTC 1295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    873 AGGAAACGCCCCCACAGGGCGCACAGCCCTACAGCCCCACACCACCCCATGGGGAGCCCTTC 932

Qy   1296 TGTGAGAGAGAGGCCTCGTTCCACCATGGCCCCTGGAAAAATGAGGATTTTAACCAAAAATT 1355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    933 TGTGAGAGAGAGGCCTCGTTCCACCATGGCCCCTGGAAAAATGAGGATTTTAACCAAAAATT 992

Qy   1356 CACACAAGCTGCTACTTTCTACAAAGACCTCACCCCAGAAGCCCAAGATCCCCTGGTCAAGAG 1415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    993 CACACAAGCTGCTACTTTCTACAAAGACCTCACCCCAGAAGCCCAAGATCCCCTGGTCAAGAG 1052

Qy   1416 AAGCTGGCCCTGATTTATGTGGAGGAAGAATCGAGCACATTTAATTGAGATCAGAGACTTATT 1475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1053 AAGCTGGCCCTGATTTATGTGGAGGAAGAATCGAGCACATTTAATTGAGATCAGAGACTTATT 1112

Qy   1476 TGGGCCCAGGCCTGGATACCCAAGAACCTCGCATAGTCATACTGCACGGGGGGCTGGAAT 1535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1113 TGGGCCCAGGCCTGGATACCCAAGAACCTCGCATAGTCATACTGCACGGGGGGCTGGAAT 1172

Qy   1536 TGGGAAGTCAACACTGGCCAGGCAGGTGAAGGAAGCCTGGGGAGGAGGCCAGCTGTATGG 1595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1173 TGGGAAGTCAACACTGGCCAGGCAGGTGAAGGAAGCCTGGGGAGGAGGCCAGCTGTATGG 1232

Qy   1596 GGACCGGCTTCCAGCAATGTCTTCTTCTACTTCAGCTGCAGGAGACTGGCCCAGTCCAAGGTGGT 1655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1233 GGACCGGCTTCCAGCATGTCTTCTTCTACTTCAGCTGCAGGAGACTGGCCCAGTCCAAGGTGGT 1292

Qy   1656 GAGTCTCGCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGGCTCCCATTAGACAGAT 1715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1293 GAGTCTCGCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGGCTCCCATTAGACAGAT 1352

Qy   1716 CCTGTCTAGGCCAGAGCGGGCTGCTCTCACCCTCGATGTGTTAGATGAGCCAGGCCAGGGGT 1775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1353 CCTGTCTAGGCCAGAGCGGGCTGCTCTCACCCTCGATGTGTTAGATGAGCCAGGCCAGGATGGGGT 1412

Qy   1776 CTTGCAGGAGGAGGCCGAGTTCTGAGCTCTCGTCACTGGAGGCATCGGCCGAGGATGC 1835
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1413 CTTGCAGGAGGACCCGAGTTCTGAGCTCTCGTCGTCTGCACTGGAGGCATCGGCCGAGGATGC 1472

Qy   1836 ACTGCTGGGGCAGTTTGCTGGGGAAAAACTATACTTCCCGGAGGCCATCCTTTTCTGATCACGGC 1895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1473 ACTGCTGGGGCAGTTTGCTGGGGAAAAACTATACTTCCCGGAGGCCATCCTTTTCTGATCACGGC 1532
```

```
Db 3693 CACATCCCTGTTCCAAGTGGCCCACTTTAAAGAGGAGGGGATGGCTTCCTGGGAGAGCCAGC     3752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4116 CAGGGTGGAGCTGCATCGATCACATAGTTCTGGAAAAACCCCTTAAAGAGGAGGTCCT          4175

Db 3753 CAGGGTGGAGCTGCATCACATAGTTCTCGAAAAACCCCAGCTTTTCCCCTTGGGAGTCCT     3812
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4176 CCTGAAAATGATCCATATAATGCCCTGCCTTCATTCCGTCACCTCTGTGGTTGTCTTTA       4235

Db 3813 CCTGAAAATGATCCATAATGCCCTGGCGCTTCATTCCGTCACCTCTGTGGGTTGTCTTTA     3872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4236 CCACCGCGTCCATCCTGAGGAAGTCACCTTCCACCTTCACCAAGTGACTGACTGCTC       4295

Db 3873 CCACCGCCTTCCATCCCTGAGGAAGTCACCTTCCACCTTCACCAAGTGACTGACTGCTC     3932
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4296 CATTCGGAAGGAACTGGGAGCTCTCGCTCTATGGGAAGCCCTTGGAGGAAGCGCCTTGGA    4355

Db 3933 CATTCGGAAGGAACTGGGAGTCTCTGCTCTATCGGAAGCCCTTGGAGGAAGACCAGCTTGGA  3992
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4356 GTTCTACGTTGGCCAACTTGGGGATCAGGGATCAGGTGCGCAAGTGAAAGAACAAGAAAGATGA 4415

Db 3993 GTTCTACGTTGGCCACTTGGGGATCAGGGATCAGGGATCAGGTGCCAAGTGAAAGAACAAGAAAGATCA  4052
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4416 GACTCTGGGTGTGGGAGGGCCTTGGTGAAAACCAGGAAGAGATCTCATGCCTGCCAACTACTCTGAT   4475

Db 4053 GACTCTGGTGTGGGAGGGCCTTGGTGAAAACCAGGAAGAGATCTCATGCCCTGCCAACTACTCTGAT    4112
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4476 CCCTTCCAGCCCGCATAGCCGTACCTTCACCTCTGGATGCCCCGCAGTTGCTGCACTTTGT     4535

Db 4113 CCCTTCCAGCCCGCATAGCCGTACCTTCACCTCTGGATGCCCCGCAGTTGCTGCACTTTGT     4172
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4536 GGACCAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGGACAA    4595

Db 4173 GGACCAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGGACAA    4232
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4596 ACTGCCATGGACAGGTGCTGCTCAGGACGGAGCCGAGTACTGCTGCTGGCTGAGACACGAG   4655

Db 4233 ACTGCCATGGACAGGTGCTGCTCGAGCCCAGGAGCAGCGTACGAGAGGGTGCTGGCTGAGAACACGAG   4292
          ||||||||||||||||||||||||||||||||||||||||||||||
Qy 4656 GCCCAGCCAGATGCGGAAGCTGTTCAGCTTGAGCCAGTGATGGCAAAGA     4715

Db 4293 GCCCAGCCAGATGCGGAACCTGTTCAGCTTGAGCCAGTGTCTGGGACCGGAAGTCAAAGA    4352
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4716 TGGACTCTACCAAGCCCTGAAAGGAGGACCCCATCCTCACCTCACCTATTATGGAACTCTGGGAA   4775

Db 4353 TGGACTCTACCAAGCCCTGAAAGGAGGAGACCCCATCCTCACCTCACCTATTATGGAACTCTGGGAGAA  4412
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4776 GGGCGAGCAAAAAGGGGACTCCTGCCACTCAGCTGACTGACTATCAACACCAGCCCCTTGACC  4835

Db 4413 GGGCGAGCAAAAAGGGGACTCCTGCCACTCAGCGAGTCTGAAGTAATGGAACTTTGATGATGCCTTTGCTGG  4472
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4836 CTTGAGTCCTGGCTTTGGCTCGTCCAGTTTCTTCTCTGCAAACAAG     4895

Db 4473 CTTGAGTCCTGGCTTTGGGCTGACCCTTCCTTGGGTCTCAGTTTCTTCTCTGCAAACAAG    4532
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4896 TTGCCATCTGGTTTGCCTTCCAGCACTAAAAGTAATGGAACTTTGATGATGCCTTTGCTGG    4955

Db 4533 TTGCCATCTGGTTTGCCTTCCAGCACTAAAAGTAATGGAACTTTGATGATGCCTTTGCTGG    4592
          ||||||||||||||||||||||||||||||||||||||||||
Qy 4956 GCAATTATTGTTCTCATTGCCAGGGGATGCCACAGGGGTGGCCTTAACACAGCA     5015

Db 4593 GCAATTATGTTCTCATTGCCAGGGATGCCACAGGGGTGGCCTTAACACAGCA          4652
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5016 TCTCAGGGGAATGTCCATCTGGAGCTGGCAAGACCCCTGCGAGACCTTCATAGAGGCCTCATCT   5075

Db 4653 TCTCAGGGAATGTCCATCTGGAGCTGGCAAGACCCCTGCGAGACCTTCATAGAGGCCTCATCT    4712
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5076 GGTGGCCACAGCAGCCAAGCCTAGAGCCCTCCGGATCCCATCCAGGGCGCAAAGAGGGAATA    5135

Db 4713 GGTGGCCACAGCAGCCAAGCCTAGAGCCCTCCGGATCCCATCCAGGGCGCAAAGAGGGAATA    4772
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5136 GGAGGGACATGGGAACCATTTGCCCTCTGGGCTGTGGTGTGTCACAGGGGGTGAGCCCCAAAAATGGGGT  5195

Db 4773 GGAGGGACATGGGAACCATTTGCCCTCTGGGCTGTGGTGTGTCACAGGGGGTGAGCCCCAAAAATGGGGT  4832
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5196 TCAGGGTGGGAGGGAGGCCCACGTGGATTCTTGTCTTTGTTACAGAAGAGATCTTACAGAAGACCAAGCC   5255

Db 4833 TCAGGGTGGGAGGGAGGCCACGTGGATTCTTGTCTTTGTTACAGAAGAGATCTTACAGAAGACC-----4892
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5256 AACAGAGTAAAAGTGGAAAGGAAGTTTATTCAGAAAATAAAGGAGTATCGCAGTCTCTTTTAG      5315

Db 4893 AACAGAGTAAAAGTGGAAAGGAAGTTTATTCAGAAAATAAAGGAGTATCACTCTGCTCTTTTAG        4952
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5316 AATTTGTCTAGGCAGGGCCTTTCCAGTTTTTTACCAGAAAAACCCCTATAAATTAAAATTTTT       5375

Db 4953 AATTTGTCTAGGCAGACTTTCCCAGTTTTTTACCAGAAAAACCCCTATAAATTAAAATTTTTT        5012
          ||||||||||||||||||||||||||||||||||||||||||
Qy 5376 ACTTAAAATTTAAGAATTAAAATCAAAAAAAAAAAATGAAAATAAAAGGAATAAGAA       5435

Db 5013 ACTTAAAATTTAAGAATTAAAATCAAAAAAAAAATGAAAATAAAAGGAATAAGAA               5072
          ||||||||||||||||||||||||||||||
Qy 5436 GTTA 5439
          ||||
Db 5073 GTTA 5076
```

```
XX
PS   Claim 5; Page 131-132; 138pp; English.
XX
CC   The present sequence is human G-protein coupled receptor-8 (GCREC-8)
CC   cDNA. The present invention relates to GCREC protein and nucleic acids
CC   encoding them. GCREC protein, its agonist or antagonist are useful for
CC   treating diseases or conditions associated with decreased expression
CC   or overexpression of functional GCREC in a patient, where the disorder
CC   is selected from cell proliferative disorders such as actinic keratosis,
CC   arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
CC   cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
CC   disease, Huntington's disease, Parkinson's disease, cardiovascular
CC   disorders such as hypertension, vasculitis, varicose veins, gastro-
CC   intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
CC   pancreatitis, autoimmune/inflammatory disorders such as acquired
CC   immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
CC   uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC   infections, trauma and metabolic disorders such as diabetes, obesity,
CC   osteoporosis. GCREC proteins and their cDNAs are used to assess the
CC   effects of exogenous compounds on the expression of GCREC sequences.
CC   GCREC cDNA is useful to create knock in humanised animals (pigs) or
CC   transgenic animals (mice or rats) to model human disease, for
CC   therapeutic or diagnostic purposes, for somatic or germline gene
CC   therapy, to generate hybridisation probes useful in mapping the
CC   naturally occurring genomic sequence, and in molecular biological
CC   techniques.
XX
SQ   Sequence 5122 BP; 1184 A; 1460 C; 1401 G; 1077 T; 0 other;
XX

Query Match        87.5%;    Score 4765.6;    DB 22;    Length 5122;
Best Local Similarity   96.9%;
Matches 4933;    Conservative    0;    Pred. No: 0;    Mismatches    24;    Indels    133;    Gaps    2;
```

```
Qy     35 GAGCCAGCCAGCCCCGGGGCGCTTCCCACTCCTGGGTTCTTCTGAAAGCCATCCCTTGCTCTCTGCGGGCTC   94
          |||||||||||||||||||| |||||  ||| |||||||||| ||||||||||||||||||||||||||||||
Db      1 GAGCCAGCCAGCCCCGGGGCTTCCCACCTCCACTCGGGTTCTTCTGAAAGCCATCCCTTGCTCTCTGAAGA   60

Qy     95 CTCCCACCCCACCCTTCTTCTTCAGCCTGCAGCTCAAGGTTGATCTCAGGACTCCAGGACC  154
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCCCACCCCACCCTTCTTCTTCAGCCTGCAGCTCAAGGTTGATCTCAGGACTCCAGGACC  120

Qy    155 CAGGAGAGGGAAGAAATCTGAGGAACACAGAACAGTGAGCGTTGCCCACACCCCATCTCCC  214
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGAGAGGGAAGAAATCTGAGGAACACAGAACAGTGAGCGTTGCCCACACCCCATCTCCC  180

Qy    215 GTCACCACACTCTCCCCCTCACCCCTCGCCTGGACCCCCATCCCCAGGAC  274
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCACCACATCTCCCCTCTGTAACACCCCTCGCCTGGACCCCCATCCCAGGAC  240

Qy    275 CTCCCTATCAGCTGACTTCTTCTTCCAGTGTGTCAGGCCCCTTCGGGCTCCTCCCTCCCCT  334
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTCCCTATCAGCTGACTTCTTCTTCCAGTGTGTCAGGCCCCTTCGGGCTCCTCCCTCCCCT  300

Qy    335 GGGCTTTTCCTACCACTCCCCCCTCTATCCGGCGTCTCTTCGGTGCCCTGGGGATTTTATA  394
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGCTTTTCCTACCACTCCCCCTCTATCCGGCGTCTCTTCGGTGCCCTGGGGATTTTATAA  360

Qy    395 AACTGGGGTTCCGGAATGCTGAAATAAGAGACGGGTAAGAGCCAAAGGCAAAAGGACACTGTT  454
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACTGGGGTTCCGGAATGCTGAAATAAGAGACGGGTAAGAGCCAAAGGCAAAAGGACACTGTT  420

Qy    455 CTCTGCCTGCCTGATACCCTCACACCTGGGAACATCCCCAGACACCCTCTTAACTCCG  514
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTCTGCCTGCCTGATACCCTCACACCTGGGAACATCCCCAGACACCCTCTTAACTCCG  480

Qy    515 GGACAGAGAATGGCTGTGGCGGGAGCCTGGGGCGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGA  574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGACAGAGAATGGCTGTGGCGGGAGCCTGGGGCGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGA  540

Qy    575 AGGAGGAGCTGAAGGAGTTCCAGCTTCTTCGCTCGGCCAATAAAAGCGCACTCCAGGAGCTCTT  634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGAGGAGCTGAAGGAGTTCCAGCTTCTTCGCTCGGCCAATAAAAGCGCACTCCAGGAGCTCTT  600
```

```
Qy    635 CGGGGTGGAGCACCCCGGCTCAGCCCAGAGAAGAAGAGGAGGTGGCATGGGAGCAGATGG  694
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGGGTGGAGCACCCCGGCTCAGCCCAGAGAAGAAGAGGAGGTGGCATGGGAGCAGATGG  660

Qy    695 TGGGCTCAGTATGGGGGAGCAGCGGGGCCTGGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG  754
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGGGCTCAGTATGGGGGAGCAGCGGGGCCTGGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG  720

Qy    755 GGCTGAAGGTCACTCTTGCGCCCCAGCCAGGAAGAGGGCAGGGCCACTCTCCTTCCTTCATTCCCT  814
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGCTGAAGGTCACTCTTGCGCCCCAGCCAGGAAGAGGGCAGGGCCACTCTCCTTCCTTCATTCCCT  780

Qy    815 ACAGCCCAAGTGAACCCACCTGGGGTCTCCCAGCCAACCCACCTCCACCGCCAGTGCTAA  874
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ACAGCCCAAGTGAACCCACCTGGGGTCTCCCAGCCAACCCACCTCCACCGCCAGTGCTAA  840

Qy    875 TGCCCTGGATCCATGAATTGCCGGGCGGGGGGTGCACCCGGGGCTCAGGAGGAAGGGTTTGA  934
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGATCCATGAATTGCCGGGCGGGGGGTGCACCCGGGGCTCAGGAGGAAGGGTTTGA  900

Qy    935 GACAGCTGCCTGCTGACACATCTGGGACCCGCGTTGGGAGAGAGAAATCTCTGCCTCACTCTCCTACC  994
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GACAGCTGCCTGCTGACACATCTGGGACCCGCGTTGGGAGAGAGAAATCTCTGCCTCACTCTCCTACC  960

Qy    995 AAGCTCTTCCAAGCTCCCCAGACCCATGAGTGCTCTCCAAGCCACCAGGAGTCACCCAACGCCCCA  1054
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AAGCTCTTCCAAGCTCCCCAGACCCATGAGTGCTCTCCAAGCCACCAGGAGTCACCCAACGCCCCA  1020

Qy   1055 CATCCCACAGCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCTACCAGCCAGAG  1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCCCACAGCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCTACCAGCCAGAG  1080

Qy   1115 AGCAGGAGGGCTTCCCTGGGGACCCAATGGCCCTCTGGGATGAACAGTCAGCGAATTTACTACACAG  1174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGCAGGAGGGCTTCCCTGGGGACCCAATGGCCCTCTGGGATGAACAGTCAGCGAATTTACTACACAG  1140

Qy   1175 AAATCAGAGAAAAGAGAGAGAGAGAGAATCAGAGAAAAGGCAGGGGCCCCCCATGGGCAGCGGGTGG  1234
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AAATCAGAGAAAAGAGAGAGAGAGAGAATCAGAGAAAAGGCAGGGGCCCCCCATGGGCAGCGGGTGG  1200

Qy   1235 TAGGAACGCCCCACAGGGCGCACCAGGCCTACAGCCCCCATGGGGGAGCCTT  1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAGGAACGCCCCACAGGGCGCACCAGGCCTACAGCCCCCATGGGGGAGCCTT  1260

Qy   1295 CTGTGAGAGAGAGGCCTCTGTTCCACATGGCCCTGGGAAAAATGAGGGATTTTAACCAAAAAAAT  1354
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTGTGAGAGAGAGGCCTCTGTTCCACATGGCCCTGGGAAAAATGAGGGATTTTAACCAAAAAAAT  1320

Qy   1355 TCACACAGCTGCTACTTCTACAAAGAACCTCACCCCAGGATCCCCTGGGTCAAGA  1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCACACAGCTGCTACTTCTACAAAGAACCTCACCCCAGGATCCCCTGGGTCAAGA  1380

Qy   1415 GAAGCTGGCCTGATTATGTGGAGGAATGAGGACACATTTAATTGAGATCAGAGACTTAT  1474
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GAAGCTGGCCTGATTATGTGGAGGAATGAGGACACATTTAATTGAGATCAGAGACTTAT  1440

Qy   1475 TTTGGCCCAGGCCTGGATACCCAGAGAACCTCGAGCAGTCATTCATCGAGCTGCTGGAA  1534
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTTGGCCCAGGCCTGGATACCCAGAGAACCTCGAGCAGTCATTCATCGAGCTGCTGGAA  1500

Qy   1535 TTGGGAAGTCAACACTGGCCAGGCAGGGTGAGGGAAGGCCTGGGGAGGCCGAGGTGTATG  1594
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTGGGAAGTCAACACTGGCCAGGCAGGGTGAGGGAAGGCCTGGGGAGGCCGAGGTGTATG  1560

Qy   1595 GGGACCGGCTTCCAGCATGGTCTTCTACTTCAGCTGCAGAGAGCTGGCCCAGGGTGG  1654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GGGACCGGCTTCCAGCATGTCTTCTACTTCAGCTGCAGAGAGCTGGCCCAGGGTGG  1620

Qy   1655 TGAGTCTCGGTTGAGTCTCATCGGAAAAAGATGGGACAGCACTCCATTAGACAGA  1714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TGAGTCTCGGTTGAGTCTCATCGGAAAAAGATGGGACAGCACTCCATTAGACAGA  1680

Qy   1715 TCCTGTCTAGGCCAGAGCGCGCTGCTCTTCATCCTGATGTGTTAGAGTGAGCCAGGAGGATGGG  1774
```

```
Qy   3935 AGATCAACCCACAGCACAGCAGCTGGATGGATGGCAGGGCGGCCTCTGCTGCTGGACATCAAGGCTGAGC 3994
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 AGATCAACCCACAGCACAGCAGCTGGATGGATGGCAGGGCGGCCTCTGCTGCTGGACATCAAGGCTGAGC 3960

Qy   3995 CTGGAGCTGTGGAAGCTGTGCACCTGCTGATGGTGGCAGGGCTTCTGCTGGGACAACGGCCCATGTGG 4054
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 CTGGAGCTGTGGAAGCTGTGCACCTGCTGATGGTGGCAGGGCTTCTGCTGGGACAACGGCCCATGTGG 4020

Qy   4055 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTGCTTCCTGGAGAAGCCAG 4114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTGCTTCCTGGAGAAGCCAG 4080

Qy   4115 CCAGGGGTGGGAGCTGCATCACATAGTTCTTGGAAAACCCCAGCTTCTCCCCCTTGGGAGTCC 4174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 CCAGGGGTGGGAGCTGCATCACATAGTTCTTGGAAAACCCCAGCTTCTCCCCCTTGGGAGTCC 4140

Qy   4175 TCCTGAAAATGATCCATAATGCCCTGCCCTTCAATCCCGTCACCTCTGTGGGTTGCTTT 4234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TCCTGAAAATGATCCATAATGCCCTGCCCTTCAATCCCGTCACCTCTGTGGGTTGCTTT 4200

Qy   4235 ACCACCCGCGTCCATCCTGAGGAGGAAGTCACCTCCTTCCACCTTCGTGTTCCCAAGTGACTGCT 4294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 ACCACCCGCGTCCATCCTGAGGAGGAAGTCACCTCCTTCCACCTTCGTGTTCCCAAGTGACTGCT 4260

Qy   4295 CCATTCGG------ 4302
          ||||||||
Db   4261 CCATTCGGAAGGCCCATAGATGATGACTAGAAAATGAAAATTCCAGTTTGTGCGAATCCACAAGC 4320

Qy   4303 ------------AAGGAACTGGAGCTCTGCTTATCTCGATCTCGAAAGCCCTGGAGGAAGAACC 4342
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CACCCCGCTGACCCCACTTTATATGGGCTGTCGTTCTACACTGTGTCTGGGGTTCAG 4380

Qy   4343 AGCTGTGTTCTCGGAGTTCTTACGTTGGGCCCACTTGGGGATCAGGGGCTGCAGCTGCAAGTGAAAG 4402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 GGATGCTGGGAAATACTCCCAAGGAACTGGAGCTCTGCTATCGAAAGCCCTGGAGGAAGAACC 4440

Qy   4403 ACAAGAAAAGATGAGACTCTGGTGGTTGGTGGGGAGGCCCTTGGTGTTGGTGAAACCAGGAGGAGATCTCATGCCTG 4462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 AGCTGGTTCTCGGAGTCCTACGTTGGCCCACTTGGGGATCAGGGGATCAGCAGTGAAAG 4500

Qy   4463 CAACTACTCTGATCCCTCCAGCCGCTACCTTCACCTTCTCTGGATGCCCCGCAGT 4522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 ACAAGAAAAGATGAGACTCTGGTGTGGTGGGGAGGCCCTTGGTGAAACCAGGAGATCTCATGCCTG 4560

Qy   4523 TGCTGCACTTTGTTGGACCAGTAGTCGAGAGCAGCAGCTGATGATAGGTGAGAGCAGCAGCAGCCGGAGTGGGACATCGGCTGGGGAGG 4582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 CAACTACTCTGATCCCTCCAGCCGCATAGCCGTTACCTTCACCTCTGGATGCCCCGCAGT 4620

Qy   4583 TTGTCTTGGACAAACTCGATGGACAGGTGCTGAGCCAGGAGCTGCAGTAGCACGTAGTAGGTGCTGG 4642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 TGCTGCACTTTGTTGGACCAGTAGTAGCAGTAGCAGTAGCAGACCCGAGTGACATGGGTGAGAGG 4680

Qy   4643 CTGAGAACACAGAGGCCCAGCCAGATGCGGAAGCTGTTCAGCTGCTTGAGCGAGTCCTGGGACC 4702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TTGTCTTGGACAAACTCGATGGACAGGTGCTGAGCCAGGAGCTGCAGTAGCACGTAGTAGGTGCTGG 4740

Qy   4703 GGAAGTGCAAAGATGGAGCTCTACCAGCGCCTCGAAGAGGAGAGGACACCCATCTCCACTCTCACCTCATTATGTG 4762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CTGAGAACACAGAGGCCCAGCCAGAAGCTGTTCAGCTTGAGCGAGTCCTGGGACC 4800

Qy   4763 AACTCTGGGAGGAGAGGCAGCAAAAAGGACTCCTGCCACTCAGCAGTGAAGTGTATCAACA 4822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GGAAGTGCAAAGATGGACTCTACCAAGCCCATCCTCACCTCACTCACCTCACCCTCCACC 4860

Qy   4823 CCAGCCCTTGACCCTTGAGTCCTGGCTTGACCCTTCTTTCAGTTTCTTT 4882
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 AACTCTCGGGAGAAGGGCAGCAAAAAAGGGACTCCTGCCACTCAGCAGTGAAGTATCAACA 4920

Qy   4883 CTCTGCAAACAAGTTGCCATCTCTGGTTTTGCCTTCCCAGCACTACTTAAAGTAATGGAACTTTGATG 4942
          ||||||||||||||||| |||||||||| |||||||||||||||||  |||||||||||||||  |||
Db   4921 CCAGCCCTTGACCCCTTGAGTCCTGGGCTTGACCCTTCTTTGGGTTCTCTCAGTTTCTTT 4980

Qy   4943 ATGCCCTTTGCTCTGGGCATTATGTGTCCAATGCCAGGGGATGCCACAGGGCGGGGGGCC 4992
          ||||||||||||||||||||||||||||||| ||||||||||||  ||||||| || |||
Db   5041 GATGCCCTTGGCTCTGGGCATTATGTGTCCA--GCCAGGGGATGCACAGGGGCGCC 5089
```

RESULT 6

AAD02760

ID   AAD02760 standard; cDNA; 4422 BP.
XX
AC   AAD02760;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NB-ARC and CARD containing protein (NAC) beta isoform cDNA.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; 88.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..4422
FT   /*tag= a
FT   /product= "Human NB-ARC and CARD containing protein (NAC)
FT   beta isoform"
FT   /note= "The translation exceptions only occur when
FT   decoding the alternate version of the protein
FT   2869..2961
FT   /transl_except= (pos:3892..3915, aa:RDAGNTPQ)"
FT   misc_feature   2869..2961
FT   /*tag= b
FT   /note= "This region is not found in human NAC gamma
FT   isoform (AAD02761) and human NAC delta isoform
FT   (AAD02762) due to the alternative mRNA splicing"
FT   misc_feature   3781..3915
FT   /*tag= c
FT   /note= "This region is not found in human NAC gamma
FT   isoform (AAD02761) due to the alternative mRNA splicing"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   P-PSDB; AAY726669, AAY727211.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
PS   Claim 4; Page 127-133; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein
CC   (NAC) beta isoform cDNA. NAC beta isoform represents the NAC splice
CC   variant in which both the splice regions are present in the translated
CC   polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC   (also referred as NB-ARC domain), a caspase-associated recruitment
CC   domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
CC   aspartyl proteases, are principal effectors of apoptosis. CARD containing

CC   NAC proteins are used for screening modulators that modulates apoptosis,
CC   cytokine production, cytokine receptor signalling and other cellular
CC   processes. NAC can act as an immunogen for the production of polyclonal
CC   and monoclonal antibodies. It can also be used to diagnose and treat
CC   inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC   pathologies such as adenocarcinomas and leukaemias.
XX
SQ   Sequence 4422 BP; 1043 A; 1224 C; 1240 G; 915 T; 0 other;

Query Match   76.2%;   Score 4148;   DB 22;   Length 4422;
Best Local Similarity   97.0%;   Pred. No. 0;
Matches 4290;   Conservative   0;   Mismatches   0;   Indels 132;   Gaps   1;

```
Qy       523 ATGGCTGGCGGGGGAGGGCCGGCCCTGGCCTGGTCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGCTGGCGGGGGAGGGCCGGCCCTGGCCTGGTCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 60

Qy       583 CTGAAGGAGTTCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGTTCCTGAAGAAGGTGAG 642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTGAAGGAGTTCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCCGGGTGAG 120

Qy       643 ACACCCCGCTTCAGCCAGAAGAAGACGAGTGGCATGGGAAGGTGGCCTTCGTTACTTGGTTGGCTCAG 702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ACACCCCGCTTCAGCCAGAAGAAGACGAGTGGCATGGAGGGTGGCCTTCGTTACTTGGTTGGCTCAG 180

Qy       703 TATGGGGACAGCAGGGACCTAGCCCTCCATACCTGGGGAGCAGATGGGGGCTGAGG 762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TATGGGGACAGCAGGGACCTAGCCCTCCATACCTGGGGAGCAGATGGGGGCTGAGG 240

Qy       763 TCACTGTGTCGCCCAAGCCCAGGCCCCCAGGAGGGCGGAGCTTCCAGAGCCCACTCTCATTCCCCTACAGCCCA 822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TCACTGTGTCGCCCAAGCCCAGGCCCCCAGGAGGGGCCCAGGCTTCCAGAGCCCACTCTCATTCCCCTACAGCCCA 300

Qy       823 AGTGAACCCCACCCTGGGGTCTCCCAGCCAACCCACTCCACCCCGCGAGTTGCTAATGCCCTCAG 882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AGTGAACCCCACCCTGGGGTCTCCCAGCCAACCCACTCCACCCCGCGAGTTGCTAATGCCCTGG 360

Qy       883 ATCCATGGAAATTGCCGCGGCGGGGTGCACCCAGGGCTCAGAGAGGAGGGTTTTGAGACAGCTG 942
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ATCCATGGAAATTGCCGCGGCGGGGTGCACCCAGGGCTCAGAGAGGAGGGTTTTGAGACAGCTG 420

Qy       943 CCTGACACATCTGGACGCGCTGGAGAGAGAAATTCTCTGCCTCACTCCTCTACCAAGGCTCTT 1002
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CCTGACACATCTGGACGCGCCGCTGGAGAGAGAAATTCTCTGCCTCACTCCTCTACCAAGCTCTT 480

Qy      1003 CCAAGCTCCCCAGACCATGAGTCTCTCCAAGCCCAGGGTCACCCGAGGTCACCAGCCCCCACATCCACA 1062
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CCAAGCTCCCCAGACCATGAGTTCTCCAAGCCCAGAGTCACCCAACGCCCCCACATCCACA 540

Qy      1063 GCAGTGCTGGGGAGCTGGGGGATCCCCACCTCAGCCTAGGCACCCAGAGCAGGAG 1122
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GCAGTGCTGGGGAGCTGGGGGATCCCCACCTCAGCCTAGGCACCCAGAGCAGGAG 600

Qy      1123 GCTCCTGGGACCCAAATGGGCCCTCTGGATGAAAAGTCAGGAGAATTTACTACACAGAAAATCAGA 1182
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GCTCCTGGGACCCAAATGGGCCCTCTGGATGAAACGTCAGGAGAATTTACTACACAGAAAATCAGA 660

Qy      1183 GAAAGAGAGAGAGAGAGAAATCAGAGAGAAAAGGCAGGCGCCCCCCATGGGCAGCGCGGTGGGTAGGGACG 1242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GAAAGAGAGAGAGAGAAATCAGAGAGAAATCAGAGGCAGGCGCCCCCCATGGGCAGCGCGGTGGGTAGGGACG 720

Qy      1243 CCCCCACAGGCGCACACCAGCCTACAGCCCCCACCACCATGGGAGGCCTTCTGTGAGA 1302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCCCCACAGGCGCACACCAGCCTACAGCCCCCACCACCATGGGAGGCCTTCTGTGAGA 780

Qy      1303 GAGAGCCCTCGTGTTCCACATGGGCCCTGTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG 1362
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GAGAGCCCTCGTGTTCCACATGGGCCCTGTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAG 840

Qy      1363 CTGCTACTTCTACAAAAGACCTCACCCCAGAAGCCAAGATCCCCTCGGTCAAGAGAGAAGCTGG 1422
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CTGCTACTTCTACAAAAGACCTCACCCCAGAAGCCAAGATCCCCTCGGTCAAGAGAGAAGCTGG 900

Qy      1423 CCTGATTTATGTTGGAGGAGGAATCAGGAGGGACATTTAATTGAGATCAGAGACTTATTGGCCCA 1482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 CCTGATTTATGTTGGAGGAGGAATCAGGAGGGACATTTAATTGAGATCAGAGACTTATTGGCCCA 960

Qy      1483 GGCCTGGGATACCCAAGAACCTCGGCAGTAGTCATACTGCAGGGGGGCTGCCTGGAAATTGGGAG 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GGCCTGGGATACCCAAGAACCTCGGCAGTAGTCATACTGCAGGGGGGCTGCCTGGAAATTGGGAG 1020

Qy      1543 TCAACACTGCCCAGCCAGGTGAAGGAAAGCTTGGGGGGAGAGGGCCACTGTTATGGGGACCGGC 1602
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TCAACACTGGCCAGCCAGGTGAAGGAAAGCTTGGGGGGAGAGGGCCACTGTTATGGGGACCGGC 1080

Qy      1603 TTCCAGCCATGTCTTCTCACTTCACTTGCACAGAGAGCTGGGCCCAGTCCAAGGTGGGGTGAGTCTC 1662
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TTCCAGCCATGTCTTCTCACTTCAGCTGCACAGAGAGCTGGGCCCAGTCCAAGGTGGGGTGAGTCTC 1140

Qy      1663 GCTGAGCTCAATCGGGAAAAAGAATGGGACAGCAGCCACTCCGGGCTTCCCAGTACAGATCCTGTCT 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GCTGAGCTCAATCGGGAAAAAGAATGGGACAGCAGCCACTCCGGGCTTCCCATTAGACAGATCCTGTCT 1200

Qy      1723 AGGGCCAGAGCGGGGCTGCTCTTCATCCTCGATGGTTGTAGAGTGAGCGTAGAGGTTCTTGCAG 1782
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 AGGGCCAGAGCGGGGCTGCTCTTCATCCTCGATGGTTGTAGAGTGAGCGTAGAGGTTCTTGCAG 1260

Qy      1783 GAGCCGGAGTTCTGGTAGGGGAAAAACTATACTTCCCGGAGGCCAGCCCAGCCGGCGGGATGCACTGCTG 1842
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 GAGCCGGAGTTCTGGTAGGGGAAAAACTATACTTCCCGGAGGCCAGCCCAGCCGGCGGGATGCACTGCTG 1320

Qy      1843 GGCAGTTGGCTGGGGAAAAACTATACTTCCCGGAGGCATCCTTCCTGATCAGGGCTGGACC 1902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 GGCAGTTTGGCTGGGGAAAAAGATGGGACAGCAGCCATCCTTCCTGATCAGGGCTGGACC 1380

Qy      1903 ACAGCTCTGCAGAACCTTCATTCCTTCTTTGGGAGCAGGCACGTTGGGTAGAGGTTCTTGGGG 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 ACAGCTCTGCAGAACCTCATTCCTTCTTTGGGAGCAGGCACGTTGGGTAGAGGTTCCTGGGG 1440

Qy      1963 TTCTCTGAGTCCAGCAGGAGGGAATAATTTCTACAGATAGTATTTCACAGATGAAAAGGCAAGCA 2022
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 TTCTCTGAGTCCAGCAGGAGGAATAATTTCTACAGATAGTATTTCACAGATGAAAAGGCAAGCA 1500

Qy      2023 ATTAGAGCCTTTAGGTTGGTTGGTCAAATCAAACAAAAGAGAGCTCTGGGCCCTGTGTCTGTGCCC 2082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 ATTAGAGCCTTTAGGTTGGTCAAATCAAACAAAAGAGAGCTCTGGGCCCTGTGTCTGTGCCC 1560

Qy      2083 TGGGTGTCTTGGGCTGGCCTTGCCCTGATGCACGATGAAGCGGAGGGAAGAAAAACTC 2142
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 TGGGTGTCTTGGGCTGGCCTTGCCCTGATGCACGATGAAGCGGAGGGAAGAAAAACTC 1620

Qy      2143 ACACTGACTTCCAAGACCACCACCAACCCTCTGTCCTACATTACCTTGCCCAGGCTCTCCAA 2202
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 ACACTGACTTCCAAGACCACCACCAACCCTCTGTCCTACATTACCTTGCCCAGGCTCTCCAA 1680

Qy      2203 GCTCAGCCATTGGGACCCCAGCTCAGAGACCTCTGCTCTCTCTGGCTGCTGAGGGCATCTGG 2262
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GCTCAGCCATTGGGACCCCAGCTCAGAGACCTCTGCTCTCTCTGGCTGCTGAGGGCATCTGG 1740

Qy      2263 CAAAAAAAAGAACCCTTTCAGTCCAGATGACCTCAGGAAGCAATGGGTTAGATGGGCCATC 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 CAAAAAAAGAACCCTTTCAGTCCAGATGACCTCAGGAAGCAATGGGTTAGATGGGCCATC 1800

Qy      2323 ATCTCCACCTTCCTTGAAGATGGGTTATTCTTCAAGCAGCACCCCATCCCTCTGAGCTACAGC 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 ATCTCCACCTTCTTCTTGAAAGATGGGTTATTCTTCAAGCAGCACCCCATCCCTCTGAGCTACAGC 1860

Qy      2383 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCAGCAAATGTCCTATGTCTTGGAGGATGAG 2442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCAGCAAATGTCCTATGTCTTGGAGGATGAG 1920

Qy      2443 AAGGGGGAGGAGGGTAAACATTCTAATTGCATCATAGATTTGGAAAAAGACGCGTAGAAGCATAT 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 AAGGGGGAGGAGGGTAAACATTCTAATTGCATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 1980

Qy      2503 GGAATACAATGGCCTCGTTTGGGGGCATCAACCACACACGTTTCCTATTGGGCCTGTTAAGTGAT 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2503 GGAATACAATGGCCTCGTTTGGGGGCATCAACCACACACGTTTCCTATTGGGCCTGTTAAGTGAT 2562
```

```
Qy   4591 GACAAAACTGCATTGACAGTGTGCTCAGCGAGGAGTAGGAGGAGGGTGCTGAGGAAC 4650
          ||||| |||||||||| |||| ||||||||| |||||||||||||||||||||||||
Db   4201 GACAAACTGCATTGACACAGGTGCTGCTCAGGAGAGTAGGAGGAGGGTGCTGAGGAAC 4260

Qy   4651 ACGAGGGCCCAGCTCAGATGCGGGAGAGCTGTTCAGTTGAGCCGAGTACCAGGAAGTGC 4710
          ||||||||||||  ||||||||||| ||||||||| ||||||||||||||||| |||||
Db   4261 ACGAGGGCCCAGATGCGGGAAGCTGTTCAGCTTGAGCCGAGTACCAGGGAAGTGC 4320

Qy   4711 AAAGATGGACTCTACCAAGCCCTGAAGGAGAGACCCATTCCTCACCTCATTATTGGAACTCTCGG 4770
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   4321 AAAGATGGACTCTACCAAGCCCTGAAGGAGAGACCCATTCCTTCACCTCATTATTGGAACTCTCGG 4380

Qy   4771 GAGAAGGGCAGCAAAAAGGGAGGACTCCTGCCCACTCACGCAGCTGA 4812
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 GAGAAGGGCAGCAAAAAGGGAGGACTCCTGCCCACTCACGCAGCTGA 4422
```

RESULT 7
AAD02761

ID   AAD02761 standard; cDNA; 4194 BP.
XX
AC   AAD02761;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NB-ARC and CARD containing protein (NAC) gamma isoform cDNA.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
Key CDS   1..4194
FT   /*tag= a
FT   /product= "Human NB-ARC and CARD containing protein (NAC)
FT   gamma isoform"
FT   misc_feature   1..2868
FT   /*tag= b
FT   /note= "Corresponds to 1-2868 residues of human NAC
FT   beta isoform (AAD02760)"
FT   misc_feature   2869..3687
FT   /*tag= c
FT   /note= "Corresponds to 2962-3780 residues of human NAC
FT   beta isoform (AAD02760)"
FT   misc_feature   3688..4194
FT   /*tag= d
FT   /note= "Corresponds to 3916-4422 residues of human NAC
FT   beta isoform (AAD02760)"
XX
PN   WO20016170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
DR   WPI; 2001-183258/18.
DR   P-PSDB; AAY72670.
XX

Claim 4; Page 138-143; 184pp; English.

PS   The present sequence is a human NB-ARC and CARD containing protein
CC   (NAC) gamma isoform cDNA. NAC gamma isoform represents the NAC splice
CC   variant in which both the splice regions are absent in the translated
CC   polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC   (also referred as a NB-ARC domain), a caspase-associated recruitment
CC   domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
CC   aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC   NAC proteins are used for screening modulators that modulates apoptosis,
CC   cytokine production, cytokine receptor signalling and other cellular
CC   processes. NAC can act as an immunogen for the production of polyclonal
CC   and monoclonal antibodies. It can also be used to diagnose and treat
CC   inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC   pathologies such as adenocarcinomas and leukaemias.
XX
SQ   Sequence 4194 BP; 984 A; 1167 C; 1176 G; 867 T; 0 other;

```
Query Match             74.9%;  Score 4078;  DB 22;  Length 4194;
Best Local Similarity   97.8%;
Matches  4194;          Conservative    0;  Mismatches    0;  Indels   96;  Gaps   2;

Qy    523 ATGGCTGGGCGGGAGCCTGGGCCGCCTGGGCCTGGCCTGTTGCTGTTACTTGAGGTTCCTGAAGAAGGAGGAG 582
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGGCGGGAGCCTGGGCCGCCTGGGCCTGGCCTGTTACTTGAGGTTCCTGAAGAAGGAGGAG 60

Qy    583 CTGAAGGAGTTCCAGCTTCTGCTTCGGCCCAATAAAGGCCACTCCAGGAGCTCTTCGGGTGAG 642
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| 
Db     61 CTGAAGGAGTTCCAGCTTCTCGCTTCGGCCAATAAAGGCCACTCCAGGAGCTCTTCGGGTGAG 120

Qy    643 ACACCCGCTCAGCCAGAGAGGAGGAGTGGCATGGAGGTGGCCTTCGTTCGTTCGGTGGCTCAG 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    121 ACACCCGCTCAGCCAGAGAGGAGGAGTGGCATGGAGGTGGCCTTCGTTCGTTCGGTGGCTCAG 180

Qy    703 TATGGGGAGGCAGCGGGGCGCGGGCCTGGGGACCTAGCCCTTCCATACCTTGGGGAGGGCTGAGG 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    181 TATGGGGAGGCAGCGGGGCGCGGGCCTGGGGACCTAGCCCTTCCATACCTTGGGGAGGGCTGAGG 240

Qy    763 TCACTGTGTGGCCCAGCCCCAGGAAGGGGCAGGCCACTCTCCCTCATTCCCTACAGCCCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    241 TCACTGTGTGGCCCAGCCCCAGGAAGGGGCAGGCCACTCTCCCTCATTCCCTACAGCCCA 300

Qy    823 AGTGAACCCCACCCTGGGGTCTCCCAGCCAACCACCTCCACCGCAGTGCTAAATGCCCTGG 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    301 AGTGAACCCCACCCTGGGGTCTCCCAGCCAACCACCTCCACCGCAGTGCTAAATGCCCTGG 360

Qy    883 ATCCATGAATTGCCGGGCGGGGTGCACCCCAGGGGTCAGAGAGGAGAGGGTTTTGAGACAGCTG 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    361 ATCCATGAATTGCCGGGCGGGGTGCACCCCAGGGGTCAGAGAGGAGAGGGTTTTGAGACAGCTG 420

Qy    943 CCTGACACATCTGGACGCGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCCTTACCAAGCTCTT 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    421 CCTGACACATCTGGACGCGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCCTTACCAAGCTCTT 480

Qy   1003 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGGCCCCCACCACCACA 1062
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    481 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGGCCCCCACCACCACA 540

Qy   1063 GCAGTGCTGGGGGAGCTGGGGGAATCCCCACCTCAGCCCAGCCTAGCCACCCAGAGAGCAGGAG 1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    541 GCAGTGCTGGGGGAGCTGGGGGATCCCCACCTTCAGCCCAGCCTAGCACCCAGAGAGCAGGAG 600

Qy   1123 GCTCCTGGGACCCAATGGCCCTCTGGATGAAAACGTCAGGAATTTACTTACACAGAAATCAGA 1182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    601 GCTCCTGGGACCCAATGGCCCTCTGGATGAAAACGTCAGGAATTTACTTACACAGAAATCAGA 660

Qy   1183 GAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGGCCCCCCCGGTGGTAGGAACG 1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    661 GAAAGAGAGAGAGAGAAATCAGAGAAAAGGGCAGGCCCCCCATGGGCAGCGGTGGTAGGAACG 720

Qy   1243 CCCCCACAGCGCGCACACCAGCCTACAGCCCCACCACACCACCACCATGGGAGGCCTTCTGTGAGA 1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    721 CCCCCACAGCGCGCACACCAGCCTACAGCCCCACCACACCACCATGGGAGCCTTCTGTGAGA 780
```

Qy   1303 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTAACCAAAAATTCACACAG 1362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTAACCAAAAATTCACACAG   840

Qy   1363 CTGCTACTTTCTACAAAAGACCTCACCCCAGAGCCCAGAAGAAGCTGG 1422
Db   841 CTGCTACTTTCTACAAAAGACCTCACCCCAGAAAGCCCAGAAGAAGCTGG 900

Qy   1423 CCTGATTTATGTGGAGGAGAATCAGAGACAGATTTAATTGGGCCCA 1482
Db   901 CCTGATTTATGTGGAGGAGAATCAGAGACAGATTTAATTGGGCCCA 960

Qy   1483 GGGCCTGGATACCCAAGAACCTCGGCATAGTCATACTGCAGGGGGCTGTCTGGAAG 1542
Db   961 GGGCCTGGATACCCAAGAACCTGGCATAGTCATACTGCAGGGGGAAG 1020

Qy   1543 TCAACACTGGCCAGGCAGGTGAAGGAACCTGAGAAGCCTGGGGGAGAGGCCCC 1602
Db   1021 TCAACACTGGCCAGGCAGGTGAAGGAACCTGAGAAGCCTGGGGGAGAGGCCCC 1080

Qy   1603 TTCCAGCATGTCTTCTCTACTTCAGCTGCAGAGAGTGGCCCAGGTGGTTGAGTCTC 1662
Db   1081 TTCCAGCATGTCTTCTCTACTTCAGCTGCAGAGAGTGGCCCAGGTGGTTGAGTCTC 1140

Qy   1663 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCT 1722
Db   1141 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCT 1200

Qy   1723 AGGCCAGAGCGGCTGCTGCTCGTTCTGCATGGTGTAGATGAGCGGAGGTCTTCAG 1782
Db   1201 AGGCCAGAGCGGCTGCTGCTCGTTCTGCATGGTGTAGATGAGCGGAGGTCTTCAG 1260

Qy   1783 GAGCCCGAGTTCTGAGCTCTGTGTCTGCACTGGAGCCACTGGAGATCACACGGCC 1842
Db   1261 GAGCCCGAGTTCTGAGCTCTGTGTCTGCACTGGAGCCACTGGAGATCACACGGCC 1320

Qy   1843 GGCAGTTTGCTGGGGAAACTATACTTCCCGAGGCCATCCTTCCTGATCACGGCTCGGACC 1902
Db   1321 GGCAGTTTGCTGGGGAAACTATACTTCCCGAGGCCATCCTTCCTGATCACGGCTCGGACC 1380

Qy   1903 ACAGCTCTGCAGAACCTCATTCCTTCCTTCTTTGGGAGCAGGAGGTTCCTGGGG 1962
Db   1381 ACAGCTCTGCAGAACCTCATTCCTTCCTTCTTTGGGAGCAGGAGGTTCCTGGGG 1440

Qy   1963 TTCTCTCTGAGTCCAGCAGAAGGAATATTTCACAGATATTTCACAGATGAAAAGGCAAGCA 2022
Db   1441 TTCTCTGAGTCCAGCAGAAGGAATATTTCACAGATATTTCACAGATGAAAAGGCAAGCA 1500

Qy   2023 ATTAGAGACCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTCGGGCCCTGGTGCCC 2082
Db   1501 ATTAGAGACCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTCGGGCCCTGGTGCCC 1560

Qy   2083 TGGGTGTCCTGGCTGGCCTGCACTTGCCTGATGGCAGCAGAATGAAGCGGAAGAAAACTC 2142
Db   1561 TGGGTGTCCTGGCTGGCCTGCACTTGCCTGATGGCAGCAGAATGAAGCGGAAGAAAACTC 1620

Qy   2143 ACACTGACTTCCAAGACCACCACAACAACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAA 2202
Db   1621 ACACTGACTTCCAAGACCACCACAACAACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAA 1680

Qy   2203 GCTCAGCCAATTGGGACCCCAGCTCAGAGAGCCTCTGCTCTCTGCTGCTGAGGGCCATCTGG 2262
Db   1681 GCTCAGCCAATTGGGACCCCAGCTCAGAGAGCCTCTGCTCTCTGCTGCTGAGGGCCATCTGG 1740

Qy   2263 CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGGAGCATGGGGTTAGATGGGGCCATC 2322
Db   1741 CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGGAGCATGGGGTTAGATGGGGCCATC 1800

Qy   2323 ATCTTCCACCTTCTTGACAGATGGGGTATTCTTCAAGAGATGGGTATTCTTCTGACTACAGC 2382
Db   1801 ATCTTCCACCTTCTTGACAGATGGGGTATTCTTCAAGAGATGGGTATTCTTCTGACTACAGC 1860

Qy   2383 TTCATTCACCTCTGTTTTCCAAGAGTTCTTTGCAGCCAATGTCCTATGTCGTCTTGGGAGGATGAG 2442
Db   1861 TTCATTCACCTCTGTTTTCCAAGAGTTCTTTGCAGCCAATGTCCTATGTCGTCTTGGGAGGATGAG 1920

Qy   2443 AAGGGAGGAGGGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGGTAGAAGCATAT 2502
Db   1921 AAGGGAGGAGGGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGGTAGAAGCATAT 1980

Qy   2503 GGAATACAATGGCCTGTTTTGGGGGCATCAACCACACAGGTTTCTCTGTTAAGTGGAT 2562
Db   1981 GGAATACAATGGCCTGTTTTGGGGGCATCAACCACACAGGTTTCTCTGTTAAGTGGAT 2040

Qy   2563 GAGGGGGGAGAGAGGAGGATGGAGAACATCTTTCACTGCCCGGCGGCTCGGGGGAGGAACCTG 2622
Db   2041 GAGGGGGGAGAGAGGAGGATGGAGAACATCTTTCACTGCCCGGCGGCTCGGGGGAGGAACCTG 2100

Qy   2623 ATGCAGTGGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGTCCTATTCGGAAGTCCCTCCAC 2682
Db   2101 ATGCAGTGGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGTCCTATTCGGAAGTCCCTCCAC 2160

Qy   2683 TGCTTGTGTACGGAGACTCGGAACAAAAGTTCCTGACAAGTGACTGATGGGCCCATTTCGGAAGAA 2742
Db   2161 TGCTTGTGTACGGAGACTCGGAACAAAAGTTCCTGACAAGTGACTGATGGGCCCATTTCGGAAGAA 2220

Qy   2743 CCCACCATGGTAGTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2802
Db   2221 ATGGGCAGTGGGTCGTGTAGAAAACAGAACATGGAGCGTCTTAGTGTGCACTTTCTGCATTAAATTC 2280

Qy   2803 AGCCGCCAGTGAAGAAGCTTCAGCTGCTGATTGAGGGGCAGCAGCACAGAATCAACATGGAGC 2862
Db   2281 AGCCGCCAGTGAAGAAGCTTCAGCTGCTGATTGAGGGGCAGCAGCACAGAATCAACATGGAGC 2340

Qy   2863 CCCACCATGGTAGTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2922
Db   2341 TGCTTGTGTACGGAGACTCGGAACAAAAGTTCCTGACAAGTGACTGATGGGCCCATTTCGGAAGAA 2400

Qy   2923 TTCTCCCGGTCCTCAAGGTCACCCAGAAACCTGAAGGAAGGAGCTGGAACTCGCTG 2982
Db   2401 TTCTCCGGTCCTCAAGGTCACCCAGAAACCTGAAGGAAGGAGCTGGAACTCGCTG 2460

Qy   2983 AGCCACTCTGCAGTGAAGAGTCTTTGTTAAGACCCTGAGACGGCCCTGCGCTGCCTCCTGGAG 3042
Db   2461 AGCCACTCTGCAGTGAAGAGTCTTTGTTAAGACCCTGAGACGGCCCTGCGCTGCCTCCTGGAG 2520

Qy   3043 ACCCTGCGGTTGGCTGGCTGTGCCCTCACAGCGTGCCAAGGACCTTGCCTTTGGG 3102
Db   2521 ACCCTGCGGTTGGCTGGCTGTGCCCTCACAGCGTGCCAAGGACCTTGCCTTTGGG 2580

Qy   3103 CTGGAGAGCCAAACCAGACCCTGACCGAGACTGGGACTTCAATGTGCTTCACGGATGCT 3162
Db   2581 CTGAGAGCCAAACCAGACCCTGACCGAGACTTCAATGTGCTCACCGGATGCT 2640

Qy   3163 GGAGCCAAACACCTTTGCCAGAGACTGAGCAGCGAGGTCACGGACTGCAG 3222
Db   2641 GGAGCCAAACACCTTTGCCAGAGACTGAGCAGCGAGGTCACGGACTGCAG 2700

Qy   3223 CTGGTCAGCTGTGGCCCTCAGCGTTCTGACTGTGCCTTAGTGTGCC 3282
Db   2701 CTGGTCAGCTGTGGCCCTCACGTGGCCTGACTGTGCCTTAGTGTGCC 2760

Qy   3283 AGCCCCAGCGCCTGAAGGGAGCTAGACCTGCAGCGAACAACCTGGATGACGTTGGGGTGCA 3342
Db   2761 AGCCCCAGCGCCTGAAGGGAGCTAGACCTGCAGCGAACAACACCTGGATGACGTTGGGGGCGA 2820

Qy   3343 CTGCTCGTTGAGGGGCTCAGGCATCCTCGCAAACTCATACGGCCT 3402
Db   2821 CTGCTCGTTGAGGGGCTCAGGCATCCTTCGCAAACTCATACGGCT 2867

Qy   3403 ACAACTCTGAGTGATGAGATGAGGCAGGAACTGAGGGGCCCTGGAGCAGGAGAAACCTCAG 3462
Db   3403 ------------------------ 2868

Qy   3463 CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTGAGGGGCCTGGGATACG 3522
Db   3463 ------------------------ 2868

```
Db    2868 ------------------GAAACCAAAGTGTGATGATGACCCCTACTGAGGGCCTGGATACG
                             ||||||||||||||||||||||||||||||||||||||||||||
Qy    3523 GGAGAGGATGAGTAATAGCACATCCTCACTCACTCAAGCGGCAGACTCGGATCAGAGAGGCG 3582
Qy    2908 GGAGAGGATGAGTAATAGCACATCCTCACTCACTCAAGCGGCAGACTCGGATCAGAGAGGCG 2967
...
```

RESULT 8
AAD02762
ID   AAD02762 standard; cDNA; 4329 BP.
XX
AC   AAD02762;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NB-ARC and CARD containing protein (NAC) delta isoform cDNA.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cysteine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..4329
FT   /product= "Human NB-ARC and CARD containing protein (NAC)
FT   delta isoform"
FT   misc_feature   1..2868
FT   /tag= b
FT   /note= "Corresponds to 1-2868 residues of human NAC
FT   beta isoform (AAD02760)"
FT   misc_feature   2869..4329
FT   /tag= c
FT   /note= "Corresponds to 2962-4422 residues of human NAC
FT   beta isoform (AAD02760)"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   P-PSDB; AAY72671.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
PS   Claim 4; Page 148-154; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein

CC (NAC) delta isoform cDNA. NAC delta isoform represents the NAC splice
CC variant in which one of the splice region is absent in the translated
CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC (also referred as NB-ARC domain), a caspase-associated recruitment
CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC NAC proteins are used for screening modulators that modulates apoptosis,
CC cytokine production, cytokine receptor signalling and other cellular
CC processes. NAC can act as an immunogen for the production of polyclonal
CC and monoclonal antibodies. It can also be used to diagnose and treat
CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC pathologies such as adenocarcinomas and leukaemias.
XX
SQ Sequence 4329 BP; 1018 A; 1202 C; 1208 G; 901 T; 0 other;

Query Match   72.6%;   Score 3952;   DB 22;   Length 4329;
Best Local Similarity   94.9%;
Matches 4197;   Conservative   0;   Mismatches   0;   Indels 225;   Gaps   2;

```
Qy      523 ATGGCTGGCGGAGACCCTGGGGGCCCCCCCCCCCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGAGGAGGAG  582
            ||||||||| ||| |||||||||| |||||||  |||||||||||||||||||||||||||||||||||||||
Db        1 ATGGCTGGCGGAGCCCTGGGCTTCTGTCTGCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGGTTGAG      60

Qy      583 CTGAAGGAGTTCCAGCTTCGTCTGCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGGTTGAG         642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTGAAGGAGTTCCAGCTTCGTCTGCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGGTTGAG         120

Qy      643 ACACCCCGCTCAGCCAGGAGGAGGAGTGGCATGGAGGTGGGCCTTCGTACCTGGTGGGCTCAG          702
Db      121 ACACCCCGCTCAGCCAGGAGGAGGACGAGTGGGCATGGAGAGTGGGCCTTCGTACCTGGTGGGCTCAG       180

Qy      703 TATGGGGAGCAGCGGGGGCCCTGGGACCTAGCCCTTCCATACCTGGGGAGCAGAGGGGCTGAGG         762
Db      181 TATGGGGGAGCAGCGGGGGCCCTGGGACCTAGCCCTCCATACCTGGGGAGCAGAGGGGCTGAGG        240

Qy      763 TCACTGTGCGGCCCAAGCCCAGGAAGGGGAGGGGCCAGCGCATGTCGGTGCTGGAGG          822
Db      241 TCACTGTGCGGCCCAAGCCCAGGAAGGGGAGGGGCCAGCGCACTCTCCCTCCCTCATTCCCTACCTACAGCCCA      300

Qy      823 AGTGAACCCCACCACTTCCACCGCAGTGCTAATGCCCTGG          882
Db      301 AGTGAACCCCACTGGGGTCTCCCAGCCACCTCCACCACACTTCCACGCAGTGCTAATGCCCTGG       360

Qy      883 ATCCATGAATTGCCGGCGGGGTCACCCAGGGGCTCAGAGAGAAGGGGTTTGAGACAGCAGTG          942
Db      361 ATCCATGAATTGCCGGCGGGGTCACCCAGGGGCTCAGAGAGGAAGGGGTTTGAGACAGCAGTG        420

Qy      943 CCTGACACATCTGGACGCGCGTGGAGAGAGAAATCTCTGCCTCACTCCTCCTTCCTCAGCCTTCT          1002
Db      421 CCTGACACATCTGGACGCGCGTGGAGAGAGAAATCTCTGCCTCACTCCTCCTTCCTCAGCCTTCT        480

Qy     1003 CCAAGCTTCCCCAGCACCATGAGAGTCTCAAGCAACGCGCCCCCACACTCCACA          1062
Db      481 CCAAGCTTCCCCAGCACCATGAGAGTCTCAAGCAACGCGCCCCCACACTCCACA        540

Qy     1063 GCAGTGCTGTGGGAGGCTGGGGAGATCCCCACTCAGCCCAGCCTAGCCACCCCAGAGAGCAGGAG          1122
Db      541 GCAGTGCTGGGGAGGCTGGGGGGATCCCCACTCAGCCCAGCCTAGCCACCCCAGAGAGCAGGAG        600

Qy     1123 GCTTCCTCGGGGACCCAAATGGCCTCTCGGATGAAACGTCAGGAGAATTTACTACACAGAAAATCAGA          1182
Db      601 GCTTCCTCGGGGACCCAAATGGCCTCTCGGATGAAACGTCAGGAGAATTTACTACACAGAAAATCAGA        660

Qy     1183 GAAAGAGAGAGAGGAGAAATCAGAGAAAGGCAGACGCCCCATGGGCAGCCGTGGTAGGAACG          1242
Db      661 GAAAGAGAGAGAGGAGAAATCAGAGAAAGGCAGACGCCCCATGGGCAGCCGTGGTAGGAACG        720

Qy     1243 CCCCACCAGGGCGCCACCAGCCTACAGCCCACCACCACCACCAGCCTTCTGTGTAGGA          1302
Db      721 CCCCACAGGGCGCCACCAGCCTACAGCCCACCACCACCACCAGCCTTCTGTGTAGGA        780

Qy     1303 GAGAGCCTCTGTTCCACATGGGCCCTGGGAAAAAATGAGGATTTTAACCAAAAAATTCACACAG          1362
Db     1801 TTCATTCACCTCTGTTTCCAAGAGAGTTCTTTGCAGCAAGTGTCCTATGTCTTGGGAGGATGAG        1920
```

Db   3988 CTGATCCTTCCAGCCCGCAATAGCCGTACCTTCACCTCTGGATGCCCGCCAGTTGCTGCAC 4047
Qy |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   4531 TTTGTGGACCAGTATCGAGAGCAGCAGCTGCCGGAGTGACATGGTGGTGGAGAGGTTGTCTTG 4590
Db   4048 TTTGTGGACCAGTATCGAGAGCAGCAGCTGCCGGAGTGACATCGGTGGTGGAGAGGTTGTCTTG 4107

Qy   4591 GACAAACTGCATGGAGGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGCTGGGCTGAGAAC 4650
Db   4108 GACAAACTGCATGGACGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGCTGGGCTGAGAAC 4167

Qy   4651 ACGAGGGCCCAGCCAGATGCGGAAGCTGTTCAGCTGAGCCAGTACGAGAGGGTGCTGGGC 4710
Db   4168 ACGAGGGCCCAGCCAGATGCGGAAGCTGTTCAGCTGAGCCAGTACGAGAGGGTGCTGGGC 4227

Qy   4711 AAAGATGGACTCTACCAAGCCCCTGAAAGGAGACCCATCCTCACCTCATTATGGAACTCTGG 4770
Db   4228 AAAGATGGACTCTACCAAGCCCCTGAAAGGAGACCCATCCTCACCTCATTATGGAACTCTGG 4287

Qy   4771 GAGAGAAGGGGCAGCAAAAAGGGACTCCTGCCACTCAGCAGCTGA 4812
Db   4288 GAGAGAAGGGGCAGCAAAAAGGGACTCCTGCCACTCAGCAGCTGA 4329

RESULT 9
AC   AAD02764
ID   AAD02764 standard; cDNA; 4556 BP.
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NAC beta isoform-CARD-X1 chimeric cDNA.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW   chimeric protein; 88.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..4365
FT   /*tag= a
FT   /product= "Human NAC beta isoform-CARD-X1 chimera"
FT   misc_feature   1..3234
FT   /*tag= b
FT   /note= "Corresponds to 1-3234 residues of human
FT   NAC beta isoform (AAD02760)"
FT   misc_feature   3235..4362
FT   /*tag= c
FT   /note= "Corresponds to 166-1293 residues of human
FT   CARD-X (AAD02763)"
XX
SQ   Sequence 4556 BP; 1100 A; 1223 C; 1256 G; 977 T; 0 other;

Query Match 62.5%; Score 3400.4; DB 22; Length 4556;
Best Local Similarity 93.7%;
Matches 3555; Conservative 0; Mismatches 236; Indels 3; Gaps 1;

Qy   523 ATGGCTGGGCGGAGCCTGGGCGCGCCTGGGCCTGGTTACTTGGAGTTCCTGAAGAAGGAGGAG 582
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCTGGGCGGAGCCTGGGCGCGCCTGGGCCTGGTTACTTGGAGTTCCTGAAGAAGGAGGAG 60

Qy   583 CTGAAGGAGTTCCAGCTTCGCTCCGCCAATAAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGAAGGAGTTCCAGCTTCGCTCCGCCAATAAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120

Qy   643 ACACCCGCTCAGACAGAGAACAGAGTGGCATGGAGATGGCGCCTCGTACCTGTGGCTCAG 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACACCCGCTCAGACAGAGAACAGAGTGGCATGGAGATGGCGCCTCGTACCTGTGGCTCAG 180

Qy   703 TATGGGGAGCAGCGCGGGCCTGGGAACCTAGCCTCCATACCTGGGACAGATGGGCTGAGG 762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TATGGGGAGCAGCGCGGGCCTGGGAACCTAGCCTCCATACCTGGGACAGATGGGCTGAGG 240

Qy   763 TCACTGTGCGCCAAGCCAGCAAGGAAGGGCAGGCCACTCTCCCTCATTCCCCTACAGCCCA 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCACTGTGCGCCAAGCCAGCAAGGAAGGGCAGGCCACTCTCCCTCATTCCCCTACAGCCCA 300

Qy   823 AGTGAAACCCCACCCTGGGGGTCTCCCAGCCAAACCCACCCTCCACGCGAGTGCTAAATGCCCTGG 882
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGTGAAACCCCACCCTGGGGGTCTCCCAGCCAAACCCACCCTCCACGCGAGTGCTAAATGCCCTGG 360

Qy   883 ATCCATGAATTGCCCGGCGGGGTGCACCCAGGGCTCAGAGAGGAAGGGTTTTGAGACAGCTG 942
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATCCATGAATTGCCCGGCGGGGTGCACCCAGGGCTCAGAGAGGAAGGGTTTTGAGACAGCTG 420

Qy   943 CCTGACACATCTGGACGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCCTCTACCAAGCTCTT 1002
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCTGACACATCTGGACGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCCTCTACCAAGCTCTT 480

Qy   1003 CCAAGCTCCCCAGACCCATGAGTCTCCAAGCCAGCCAGAGTCACCCAACACGACCCCATCCACA 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCAAGCTCCCCAGACCCATGAGTCTCCAAGCCAGCCAGAGTCACCCAACACGACCCCATCCACA 540

Qy   1063 GCAGTGCTGGGGAGCTGGGGGATCCCCACCTTCAGCCCAGCTTAGCCCAGAGAGCCAGGAG 1122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCAGTGCTGGGGAGCTGGGGGATCCCCACCTTCAGCCCAGCTTAGCCCAGAGAGCCAGGAG 600

Qy   1123 GCTCCTGGGACCCAATGGCCCTCTCGGATGAAAGCGTCAGGAATTTACTCACACAGAAATCAGA 1182
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCTGGGACCCAATGGCCCTCTCGGATGAAAGCGTCAGGAATTTACTCACACAGAAATCAGA 660

Qy   1183 GAAAGAGAGAGAGAGAGAAATCAGAGAGAAAGGCAGGGCCCCATGGGCAGCGCGGTGGGTAGGAGCG 1242
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GAAAGAGAGAGAGAGAGAAATCAGAGAGAAAGGCAGGGCCCCATGGGCAGCGCGGTGGGTAGGAACG 720

Qy   1243 CCCCCACAGGCGCACACCAGCCTACAGCCCCACCACCATGGGAGCGCTTCTGTGAGA 1302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

PT   apoptosis -
XX
PS   Disclosure; Page 162-168; 184pp; English.
XX
PT   The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC   beta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
CC   nucleotide binding (NB) domain (also referred as NB-ARC domain); a
CC   caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC   domain. CARD-X protein comprises a caspase-associated recruitment domain
CC   (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC   proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC   for screening modulators that modulates apoptosis, cytokine production,
CC   cytokine receptor signalling and other cellular processes. They can act
CC   as an immunogen for the production of polyclonal and monoclonal
CC   antibodies. They can also be used to diagnose and treat inflammatory
CC   disorders such as sepsis, fibrosis and arthritis and cancer pathologies
CC   such as adenocarcinomas and leukaemias.

PN   WO200116170-A2.
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999; 99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   P-PSDB; AAY72673.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of

```
Qy    3463 CTGCTCATCTTCAGCAGCAGAAACCAAGTGTGATGATGTGATGACCCTACTGAGGCTGGATACG 3522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 CTGCTCATCTTCAGCAGCAGAAACCAAGTGTGATGATGTGATGACCCTACTGAGGCTGGATACG 3000

Qy    3523 GGAGAGAGATGAGTAAATAGACACATCCTCACTCAGGGGGCCAGAGAGACTCGGATCAGAGAGGGCG 3582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 GGAGAGAGATGAGTAAATAGACACATCCTCACTCAGGGGGCCAGAGAGACTCGGATCAGAGAGGGCG 3060

Qy    3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTCAGACGTGGCAAGAGATCTTCCCAATT 3642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTCAGACGTGGCAAGAGATCTTCCCAATT 3120

Qy    3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTTGCCGTGCCT 3702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTTGCCGTGCCT 3180

Qy    3703 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGAGGCCTTTTGGGGACTGACGATGACTTCTGG---GTGAGATC 3762
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||       ||||||
Db    3181 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGAGGCCTTTTGGGGACTGACGATGACTTTCTG 3240

Qy    3763 GGCCCCACGGGGCCCTGTGGGCTACTAGTTGACAAAGAAAGAACTTGTTACCGAGTT 3822
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 GGGCCTGAAGGAAAATGTTGGAGTTGAGTTGATTGATTGAATAAGAGCACAAAACAGATACACGTT 3300

Qy    3823 CACTTCCCTGTAGCTGGCTCCTACCCGCTGGCCCAACACGGGGTCTTCTGCTTTGTTGTGATGAGA 3882
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3301 TGGTTCCCCACTGCTTGGCTCTGGCTATTCGTAGTGCTTGTTCAGCTGGCCTTCGGCTTCCTGGTTAAGG 3360

Qy    3883 GAAGCGGTGACCGTTGAGATTGAATTCTGTGTGTGTGGGGACCAGTTCCTGG---GTGAGATC 3939
           |||||||||||||||||||||||||||||||||||||||
Db    3361 GATGAGGTTCACAGTGACGGATTGCGTTTGGTTCCTGGGAGTCAGGCACCTGGCCCTGGGACCTG 3420

Qy    3940 AACCCACAGCACAGCTGGATGGTGGCAGGGGCCTTCTGCTGTGGAGTCACATCAAGGCTGAGGCCTGGA 3999
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3421 CAGCACCATGAACAGTGACTGGGCTGTGGGGGGCCCCTTGTTTGATGAGTCACTGCAGAGCCAGAG 3480

Qy    4000 GCTGTGGGAAGCTGTGTGCACCTCCCCTCACTTGTGTGGCTCTCCAAGGGGGGCCATGTGGACACA 4059
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3481 GAGGCGTGTCGCCGGAAATCCACCTCCCCCACTTCATCTCCCCGGTGAGGCTGAGGTGGACGTC 3540

Qy    4060 TCCCTGTTCCAAAATGGGCCCACTTTAAAGAGAGGGGAGGGGGCTCCTGGAGAAGCCAGCCACGG 4119
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 TCCTGGTTTCTCGGTTGCCCATTTTAAGGAATGAAGGGATGGTCCTGGGACATCCAGCGCCCG 3600

Qy    4120 GTGGGAGGTGCATCACATAGTTCTGGAAAAACCCCCAGCTTCTCCCCCTTGGGAGTCCTCCCTG 4179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3601 GTGGAGGCCTTTCTATGGTGTCCTGCTGTCCTGGAAAGGCCCACAGCTTCTCTCTGATGGGGCCATCCTGCTG 3660

Qy    4180 AAAATGATTCCATAAATGCCCTGCGGTCACCTCTGTGGTTGTTGCTTTACCAC 4239
           |||||||||||||||||||||||||||||
Db    3661 CGGATCGCCAGTGGGGACTCGGCCCTCTCCCATCCCATCCAAACACATTGATCTATTTAT 3720

Qy    4240 CGCGTCCATCCTGTGAGGAAGTCACCTTCCACCTTCACCTTCACCTGGATCCCAAGTGACTGCTCCATT 4299
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    3721 CACCCCGACCCCGAAGATATTAAGTTCCACTTGTTCCCAGCGAGCCGCCTTGCTA 3780

Qy    4300 CGGAAGGAAACTCGA 4313
           ||||||||||||
Db    3781 ACAAAGGCGATAGA 3794
```

RESULT 10
AAD02765
ID   AAD02765 standard; cDNA; 4466 BP.
XX
AC   AAD02765;
XX
DT   31-MAY-2001 (first entry)
XX
DE   Human NAC gamma or delta isoform-CARD-X1 chimeric cDNA.
XX
KW   Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW   caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW   cysteine aspartyl protease; apoptosis; cytokine production;
KW   cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW   fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW   chimeric protein; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..4275
FT   /tag= a
FT   /product= "Human NAC gamma/delta isoform-CARD-X1 chimera"
FT   misc_feature   1..3144
FT   /tag= b
FT   /note= "Corresponds to 1-3144 residues of human
FT   NAC gamma (AAD02761) or delta (AAD02762) isoform"
FT   misc_feature   3145..4272
FT   /tag= c
FT   /note= "Corresponds to 166-1293 residues of human
FT   CARD-X (AAD02763)"
XX
PN   WO200116170-A2.
XX
PD   08-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24152.
XX
PR   01-SEP-1999;   99US-0388221.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC;
XX
DR   WPI; 2001-183258/18.
DR   P-PSDB; AAY72674.
XX
PT   Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT   domains, used to produce polypeptides for screening for modulators of
PT   apoptosis -
XX
PS   Disclosure; Page 173-178; 184pp; English.
XX
CC   The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC   gamma or delta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
CC   nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC   caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC   domain. CARD-X protein comprises a caspase-associated recruitment domain
CC   (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC   proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC   for screening modulators that modulates apoptosis, cytokine production,
CC   cytokine receptor signalling and other cellular processes. They can act
CC   as an immunogen for the production of polyclonal and monoclonal
CC   antibodies. They can also be used to diagnose and treat inflammatory
CC   disorders such as sepsis, fibrosis and arthritis and cancer pathologies
CC   such as adenocarcinomas and leukaemias.
XX
SQ   Sequence 4466 BP; 1201 A; 1227 G; 963 T; 0 other;

```
Query Match             59.0%;  Score 3210.4;  DB 22;  Length 4466;
Best Local Similarity   91.3%;
Matches 3465; Conservative    0;  Mismatches  236;  Indels   93;  Gaps   2;

Qy     523 ATGGCTGGCGGGAGCCTGGGGCCGCCCTGGCCCTGTTACTTGGAGTTCCTGAAGAGAAGGAGAG 582
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGCTGGCGGAGCCTGGGGCGGCCCCTGGCCCTGTTACTTGGAGTTCCTGAAGAGAAGGAGAG 60

Qy     583 CTGAAGGAGTTCCAGCTTCTTGCTTCTCGCGCCAATAAAAGCGCAACTCCAGGGAGCTTCTTCGGGGTGAG 642
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGAAGGAGTTCCAGCTTCTTGCTTCTCGCCAATAAAAGCGCACTCCAGGGAGCTTCTTCGGGGTGAG 120

Qy     643 ACACCCGCTCAGCCAGAGAAGAGAGAGCGAGTGGCCATGGAGGTGGCCTCGTACCTGGTGGCTCAG 702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACACCCGCTCAGCCAGAGAAGAGACGAGTGGCATGGAGGTGGCCTCGTACCTGGTGGCTCAG 180
```

Qy   703 TATGGGGGAGCAGCGGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGGCTGAGG 762
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TATGGGGGAGCAGCGGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGGCTGAGG 240

Qy   763 TCACTGTGCGCGCCCAAGCCCAGGGGGCGGCCA... 822
Db   241 TCACTGTGCGCGCCCAAGCCCAGGGGGCGGCCA... 300

Qy   823 AGTGAACCCCACCTGGGGTCTCCCAGGGGG... 882
Db   301 AGTGAACCCCACCTGGGGTCTCCCAGGGGG... 360

[Sequence alignment continues — Qy (query) vs Db (database) DNA sequence alignment blocks]

```
Qy   2341 CCCACCAATGGTAGTTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCCACCAATGGTAGTTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2400

Qy   2401 TTCTCCGTTCCTCAAGGTCACCACCAGAAACCTGAAGAGGAGCTGGACTTAAGTGGAAACTCGGCTG 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTCTCCGTTCCTCAAGGTCACCACCAGAAACCTGAAGAGGAGCTGGACTTAAGTGGAAACTCGGCTG 2460

Qy   2461 AGCCAACTCTGCAGTGCAGAGAGTCTTTGTAAGACCCTGAGACCCTTGCCTCTCCTGGAG 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AGCCAACTCTGCAGTGCAGAGAGTCTTTGTAAGACCCTGAGACCCTTGCCTCTCCTGGAG 2520

Qy   2521 ACCCTGCGGGTTGGCTGGCCTCACAGCTGCAAGGACCTTGCCTTTGGG 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACCCTGCGGGTTGGCTGGCCTCACAGCTGCAAGGACCTTGCCTTTGGG 2580

Qy   2581 CTGAGAGCCAACACAGACCCTGACCGAGTGGACCTGAGCTTCAAATGTGCTCACGGATGCT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CTGAGAGCCAACACAGACCCTGACCGAGTGGACCTGAGCTTCAAATGTGCTCACGGATGCT 2640

Qy   2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAAGCTACAGCCGACTGCAG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAAGCTACAGCCGACTGCAG 2700

Qy   2701 CTGGTCAGCTGTGTGCCCTCACGTCTTGACTGTTGCTTGCCTTTAGTGCC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CTGGTCAGCTGTGTGCCCTCACGTCTTGACTGTTGCTTGCCTTTAGTGCC 2760

Qy   2761 AGCCCCAGCCTGAAGGAGCTAGACCTGACAGCAGAACAACCTGGATGACGTTGGCGGTGGCA 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 AGCCCCAGCCTGAAGGAGCTAGACCTGACAGCAGAACAACCTGGATGACGTTGGCGGTGGCA 2820

Qy   2821 CTGCTCTCGTGTGAGGGGGCTCAGGCATCCTGCCTGAAACTCATACGCCTCG------ 2869
          ||||||||||||||||||||||||||||||||||||
Db   2821 CTGCTCTCGTGTGAGGGGGCTCAGGCATCCTGCCTGAAACTCA------- 2869

Qy   2870 ---------------------------------------------- 2869
Db   2870 GAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG 2910

Qy   2911 GGAGAGATGAGTAATAGAGCACATCCTCACTCAAGCGGCAGAGATCTCGGATCAGAGAGGGG 2970
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2911 GGAGAGATGAGTAATAGAGCACATCCTCACTCAAGCGGCAGAGATCTCGGATCAGAGAGGGG 2970

Qy   2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTGGACCTGCAACAGAGAATCTTCCCAATT 3030
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTGGACCTGCAACAGAGAATCTTCCCAATT 3030

Qy   3031 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGGGATGCCT 3090
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3031 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGGGATGCCT 3090

Qy   3091 TCTCCTGCCTCTCAAGGGACCTTTGGGGACTGAGCCTGACCGATGACTTTCTG 3150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3091 TCTCCTGCCTCTCAAGGGACCTTTGGGGACTGAGCCTGACCGATGACTTTCTG 3150

Qy   3151 GGGCCTGAAGGAGAAAGTGGAGTTGAGTTGATTAAGAGCACAAACAGATACAGCGTT 3210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3151 GGGCCTGAAGGAGAAAGTGGAGTTGAGTTGATTAAGAGCACAAACAGATACAGCGTT 3210

Qy   3211 TGGTTCCCCACTGCTGGCTCGGCTTCCTGTAAGG 3270
          ||||||||||||||||||||||||||||||||||
Db   3211 TGGTTCCCCACTGCTGGCTCGGCTTCCTGTAAGG 3270

Qy   3271 GATGAGGTCACAGTGACGATTGCGTTTGGTTCCGGAGCATCAAGGCTGAGCCTGGA 3330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3271 GATGAGGTCACAGTGACGATTGCGTTTGGTTCCGGAGCATCAAGGCTGAGCCTGGA 3330

Qy   3331 CAGACACCATGAACACAGTGGGCTGGTGGTGGGGCGGCGCCGCCCCTTGTTGTGATGGTCACTCGAGAGCCGGAG 3390
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3331 CAGACACCATGAACACAGTGGGCTGGTGGTGGGGCGGCGCCGCCCCTTGTTGTGATGGTCACTCGAGAGCCGGAG 3390

Qy   3391 GAGGCTGGTGCCGCGGAAATCCACCTCCCCCACTTCATCTCCCTCCAAGGTGGGGGAGTC 3450
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3391 GAGGCTGGTGCCGCGGAAATCCACCTCCCCCACTTCATCTCCCTCCAAGGTGGGGGAGTC 3450

Qy   3451 TCCTGGTTTCTCGTTGCCCATTTTAAGAATGAAGGATGGTGCCGGAGCATCCACCGCCGG 3510
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3451 TCCTGGTTTCTCGTTGCCCATTTTAAGAATGAAGGATGGTGCCGGAGCATCCACCGCCGG 3510

Qy   3511 GTGGAGCGCATCACATAGTTCTCGGGAAAACCCAGCTTCTCCCCGTCTGGAGCGTCCTCTG 3570
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3511 GTGGAGCGCATCACATAGTTCTCGGGAAAACCCAGCTTCTCCCCGTCTGGAGCGTCCTCTG 3570

Qy   3571 CGGATCGCCAGTGGGGACTCCGCCCTTCTCCATCCCCAATCACTTCCAACACATTGATCTATTAT 3630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3571 CGGATCGCCAGTGGGGACTCCGCCCTTCTCCATCCCCAATCACTTCCAACACATTGATCTATTAT 3630

Qy   3631 CACCCCCACCCCGAAGATTATTAAGTTCACTTGTACCTTGTCCCAGCGACGACGCCTTGCTA 3690
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3631 CACCCCCACCCCGAAGATTATTAAGTTCACTTGTACCTTGTCCCAGCGACGACGCCTTGCTA 3690

Qy   3691 ACAAAGGCGATAGA 3704
          ||||||||||||||
Db   3691 ACAAAGGCGATAGA 3704
```

RESULT 11
AAA78392

ID   AAA78392 standard; cDNA; 2657 BP.
XX
AC   AAA78392;
XX
DT   20-NOV-2000 (first entry)
XX
DE   Human secreted protein gene 12 SEQ ID NO:22.
XX
KW   Human; secreted protein; cytostatic; antianaemic; antidiabetic;
KW   antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
KW   antipsoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
KW   neuroprotective; antimicrobial; antiparkinsonian; cancer;
KW   immune system disorder; angiogenesis; hyperproliferative disorder;
KW   cardiovascular disorder; apoptosis; neurological disease;
KW   infectious disease; wound healing; ss.
XX
OS   Homo sapiens.
XX
PN   WO200035937-A1.
XX
PD   22-JUN-2000.
XX
PF   16-DEC-1999; 99WO-US29950.
XX
PR   17-DEC-1998; 98US-0112809.
PR   18-DEC-1998; 98US-0113006.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
PI   Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
XX
DR   WPI; 2000-431566/37.
DR   P-PSDB; AAB24448.
XX
PT   Forty seven human nucleic acids encoding secreted proteins, useful in
PT   the treatment, prevention and diagnosis of cancers, disorders of the
PT   immune system, angiogenesis disorders, neurological diseases and
PT   hyperproliferative disorders -
XX
PS   Claim 1; Page 449-450; 562pp; English.
XX
CC   The polynucleotide sequence given in AAA78381 to AAA78432 encode the
CC   human secreted proteins given in AAB24437 to AAB24604. Human secreted

```
CC   proteins have activities based on the tissues and cells the genes are
CC   expressed in. Examples of activities include: cytostatic; antianaemic;
CC   antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
CC   antiarthritic; antipsoriatic; antiangiogenic; cardiant; anti-HIV;
CC   nootropic; neuroprotective; antimicrobial and antiparkinsonian.
CC   Human secreted protein polynucleotides, polypeptides, antagonists and/or
CC   agonists may be useful in treating, preventing, and/or diagnosing other
CC   diseases, disorders, and/or conditions such as: (a) cancers; (b)
CC   diseases of the immune system; (c) angiogenesis disorders; (d)
CC   hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
CC   associated with increase apoptosis; (g) neurological diseases; and
CC   (h) infectious diseases. They are also used to promote wound healing.
CC   AAA78380 and AAB24436 represent sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2657 BP; 664 A; 696 C; 726 G; 571 T; 0 other;

Query Match      47.3%;  Score 2572.6;  DB 21;  Length 2657;
Best Local Similarity   99.4%;
Matches 2614;  Conservative   0;  Mismatches  14;  Indels   3;  Gaps   3;
```

```
Qy   2814 GAAGAAGCTTCAGCTGATTGAGGCAGCAGCACAGATCAACATGGAGCCCCACCCATGGT 2873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 GAAGAAGCTTCAGCTGATTGAGGCAGCAGCACAGATCAACATGGAGCCCCACCCATGGT 68

Qy   2874 AGTCCTGTTCAGGTGGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCTCCGTCCT 2933
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 AGTCCTGTTCAGGTGGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTCTTCTCCGTCCT 128

Qy   2934 CAAGGTCACCAGAAAACCTGAAGGAGCTGACCTAAGTGGAAACTCGCTGAGCCACTCTGC 2993
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 CAAGGTCACCAGAAAACCTGAAGGAGCTGACCTAAGTGGAAACTCGCTGAGCCACTCTGC 188

Qy   2994 AGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTCGCTGCCTCTCCTGGAGACCCTGCGGTT 3053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 AGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTCGCTGCCTCTCCTGGAGACCCTGCGGTT 248

Qy   3054 GGCTGGCTGCTGTGGGCCTCACCAGCTGAAGGAGACCCTTGCCTTTGGGCTGAGGAGCCAA 3113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 GGCTGGCTGCTGTGGGCCTCACCAGCTGAAGGAGACCCTTGCCTTTGGGCTGAGGAGCCAA 308

Qy   3114 CCAGACCCCTGACCGAGCTGGAACCTGAGCTTCAATGTGCTTCACGGATGCTGGGACCAAACA 3173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CCAGACCCCTGACCGAGCTGGAACCTGAGCTTCAATGTGCTTCACGGATGCTGGGACCAAACA 368

Qy   3174 CCTTTGCCAGAGACTGACACAGCCGAGCTGCAAGCTACAGCAGCTCAGCTGGTCAGCTG 3233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 CCTTTGCCAGAGACTGACACAGCCGAGCTGCAAGCTACAGCAGCTCAGCTGGTCAGCTG 428

Qy   3234 TGGCCTCACGTCTGCTACTGTCCCAGGACCTGGCCTCCTCTGTGTGTCTTAGTGTCCAGCCT 3293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    429 TGGCCTCACGTCTGCTACTGTCCCAGGACCTGGCCTCCTCTGTGTGTCTTAGTGTCCAGCCT 488

Qy   3294 GAAGGAGCTAGACGTGCAGCAGAAACAACCTGGATGACGTTGGGCGTGCCGACTGTCTGTGA 3353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 GAAGGAGCTAGACGTGCAGCAGAAACAACCTGGATGACGTTGGGCGTGCCGACTGTCTGTGA 548

Qy   3354 GGGGCTCAGGCATCCTGCCTCGCAAACTCATACCCCTGGGCGTGCAGCAGAACTCTCAGG 3413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549 GGGGCTCA-GCATCCTGCCTCGCAAACTCATACCCCTGGGCGTGCAGCAGAACTCTCAGG 607

Qy   3414 TGATGAGATGAGGCAGGAACTGAGGCGCCCTGGGACGAGGAGAAAACCTCAGCTGCTCATCTT 3473
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 TGATGAGATGAGGCAGGAACTGAGGCGCCCTGGGACGAGGAGAAAACCTCAGCTGCTCATCTT 667

Qy   3474 CAGCAGACGGGAAACCAAGTGTGATGATGACCCCTACTGA-GGCCTGGATACCGGAGGATGAG 3533
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    668 CAGCAGACGGGAAACCAAGTGTGATGATGACCCCTACTGA-GGCCTGGATACCGGAGGATGAG 726

Qy   3534 TAATAGCACATCCTCACTTCAAGCGGCACAGAGACTCGGATCAGAGAGGGGGCCTTCCCATGT 3593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    727 TAATAGCACATCCTCACTTCAAGCGGCACAGAGACTCGGATCAGAGAGGGGGCCTTCCCATGT 786

Qy   3594 TGCTTCAGGCTAATCTCAAAACTTCCTTGGACCTGAGCAGAGACAAGATCTTCCCAATTGCT 3653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    787 TGCTTCAGGCTAATCTCAAAACTTCCTTGGACCTGAGCAGAGACAAGATCTTCCCAATTGCT 846

Qy   3654 AGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCCGCCTC 3713
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    847 AGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCCGCCTC 906

Qy   3714 TCAAGGGGGACCTGCTACTGAAAGCCCTTTGGGGACTGACGATGACTTCTGGGGGCCCACGGG 3773
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    907 TCAAGGGGGACCTGCTACTGAAAGCCCTTTGGGGACTGACGATGACTTCTGGGGGCCCACGGG 966

Qy   3774 GCCTGTGGGCTACTGAGGGTAGTTGACAAAGAAAAGAAACTTGTACCGGAGTTCACTTCCCTGT 3833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    967 GCCTGTGGGCTACTGAGGGTAGTTGACAAAGAAAAGAAACTTGTACCGGAGTTCACTTCCCTGT 1026

Qy   3834 AGCTGGCTCCTACCGCCTGGCCAACACGGGGTCTCGCCTTTGTGATTGATGAGAGAGGCGGTGAC 3893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 AGCTGGCTCCTACCGCCTGGCCAACACGGGGTCTCGCCTTTGTGATTGATGAGAGAGGCGGTGAC 1086

Qy   3894 CGTTGAGATGAATTCTGTGTGTGTGGGACCAGTTCCTGTGGGTGAGATCAACCCACAGCACAG 3953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 CGTTGAGATGAATTCTGTGTGTGTGGGACCAGTTCCTGTGGGTGAGATCAACCCACAGCACAG 1146

Qy   3954 CTGGATGGTGGCAGGGCCTTCTGCTTGGGACCATCAAGGCGTGGACTGAGCCTGGAGGCTGT 4013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 CTGGATGGTGGCAGGGCCTTCTGCTTGGGACCATCAAGGCGTGGACTGAGCCTGGAGGCTGT 1206

Qy   4014 GCACCTCCCTCACTTTGTGGCTCTTCCAAGGGGGCCCAGTGTGGACATCCCTGTTCCAAGT 4073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 GCACCTCCCTCACTTTGTGGCTCTTCCAAGGGGGCCCAGTGTGGACATCCCTGTTCCAAGT 1266

Qy   4074 GGGCCACACTTTAAAGAGGAGGGGGGATGCTCCTGGAGAAGCCAGCCAGGGTGGGACTGGCA 4133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1267 GGGCCACACTTTAAAGAGGAGGGGGGATGCTCCTGGAGAAGCCAGCCAGGGTGGGACTGGCA 1326

Qy   4134 CATAGTTTCTGGGAAAAACCCCAGCTTCTCCCCCCTGTGGGGAGTCCCTGAAAATGATCCATAA 4193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1327 CATAGTTTCTGGGAAAAACCCCAGCTTCTCCCCCCTGTGGGGAGTCCCTGAAAATGATCCATAA 1386

Qy   4194 TGCCCTGCGCTTCATTCCCGTCACCTCTGTGGTGTTGCTTTAACCACCGCGGGTCCATCCTGA 4253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1387 TGCCCTGCGCTTCATTCCCGTCACCTCTGTGGTGTTGCTTTAACCACCGCGGGTCCATCCTGA 1446

Qy   4254 GGAAGTCACCTTCACCTTCACCCTCTACTTGGTTGTGTGTATGTGTCTCCCATTCGGAAGGAAACTGGA 4313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1447 GGAAGTCACCTTCACCTTCACCCTCTACTTGGTTGTGTGTATGTGTCTCCCATTCGGAAGGAAACTGGA 1506

Qy   4314 GCTCTGCTATCGAAGCCCTGGGAGAAGACCAGCTGTTCTCGGGAGTTCTAGGTTGGCCACTTT 4373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1507 GCTCTGCTATCGAAGCCCTGGGAGAAGACCAGCTGTTCTCGGGAGTTCTAGGTTGGCCACTTT 1566

Qy   4374 GGGATCAGGGATCAGGCTGCAAGTGAAAGATGAAAGATGAGACTCGGTTGGTGTGGGAGGC 4433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1567 GGGATCAGGGATCAGGCTGCAAGTGAAAGATGAAAGATGAGACTCGGTTGGTGTGGGAGGC 1626

Qy   4434 CTTGGGTGAAACAGGAGATCTCAATGCCTGCCGCAACTACTCTGATCCCTCCAGCCCCATAGC 4493
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1627 CTTGGGTGAAACAGGAGATCTCAATGCCTGCCGCAACTACTCTGATCCCTCCAGCCCCATAGC 1686

Qy   4494 CGTACCTTCACCTCTGGATGCCCCGCAGTTGCTCCACTTTGTGGACCAGTATCGAGAGCA 4553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1687 CGTACCTTCACCTCTGGATGCCCCGCAGTTGCTCCACTTTGTGGACCAGTATCGAGAGCA 1746

Qy   4554 GCTGATAGCCCGAGTGACATCGGTGGAGAGTTGTTCTTGGACAAAACTGCATGGACAGGTGCT 4613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1747 GCTGATAGCCCGAGTGACATCGGTGGAGAGTTGTTCTTGGACAAAACTGCATGGACAGGTGCT 1806

Qy   4614 GAGCCAGGAGCAGTACGAGGGTGCTGCTGAGAACACGGAGGGCCCAGATGCGGAA 4673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1807 GAGCCAGGAGCAGTACGAGGGTGCTGCTGAGAACACGGAGGGCCCAGATGCGGAA 1866

Qy   4674 GCTGTTCAGCTTGAGCCAGTCCTGGGACCGGAGTGCAAAGAGATGGACTCTACCAAGCCCT 4733
Db
```

RESULT 12
AAL01744

```
ID   AAL01744 standard; cDNA; 578 BP.
XX
AC   AAL01744;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen cDNA SEQ ID NO: 1745.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000;  2000US-0179065.
PR   04-FEB-2000;  2000US-0180628.
PR   24-FEB-2000;  2000US-0184664.
PR   02-MAR-2000;  2000US-0186350.
PR   16-MAR-2000;  2000US-0189874.
PR   17-MAR-2000;  2000US-0190076.
PR   18-APR-2000;  2000US-0198123.
PR   19-MAY-2000;  2000US-0205515.
PR   07-JUN-2000;  2000US-0209467.
PR   28-JUN-2000;  2000US-0214886.
PR   30-JUN-2000;  2000US-0215135.
PR   07-JUL-2000;  2000US-0216647.
PR   07-JUL-2000;  2000US-0216880.
PR   11-JUL-2000;  2000US-0217487.
PR   11-JUL-2000;  2000US-0217496.
PR   14-JUL-2000;  2000US-0218290.
PR   26-JUL-2000;  2000US-0220963.
PR   26-JUL-2000;  2000US-0220964.
PR   14-AUG-2000;  2000US-0224518.
PR   14-AUG-2000;  2000US-0224519.
PR   14-AUG-2000;  2000US-0225213.
PR   14-AUG-2000;  2000US-0225214.
PR   14-AUG-2000;  2000US-0225266.
PR   14-AUG-2000;  2000US-0225267.
PR   14-AUG-2000;  2000US-0225268.
PR   14-AUG-2000;  2000US-0225270.
PR   14-AUG-2000;  2000US-0225447.
PR   14-AUG-2000;  2000US-0225757.
PR   14-AUG-2000;  2000US-0225758.
PR   14-AUG-2000;  2000US-0225759.
PR   18-AUG-2000;  2000US-0226279.
PR   22-AUG-2000;  2000US-0226681.
PR   22-AUG-2000;  2000US-0226868.
PR   22-AUG-2000;  2000US-0227182.
PR   23-AUG-2000;  2000US-0227009.
PR   30-AUG-2000;  2000US-0228924.
PR   01-SEP-2000;  2000US-0229287.
PR   01-SEP-2000;  2000US-0229343.
PR   01-SEP-2000;  2000US-0229344.
PR   01-SEP-2000;  2000US-0229345.
PR   05-SEP-2000;  2000US-0229509.
PR   05-SEP-2000;  2000US-0229513.
PR   06-SEP-2000;  2000US-0230437.
PR   06-SEP-2000;  2000US-0230438.
PR   08-SEP-2000;  2000US-0231242.
PR   08-SEP-2000;  2000US-0231243.
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232080.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232397.
PR   14-SEP-2000;  2000US-0232398.
PR   14-SEP-2000;  2000US-0232399.
PR   14-SEP-2000;  2000US-0232400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235834.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   02-OCT-2000;  2000US-0236802.
```

Sequence alignment:

```
Db   1867 GCTGGTTCAGCTTGGAGGGACCCAGTCCTGGGACCGGAAGTGCAAAGATGGAGAGACTTCTACCAAGCCCT 1926
          |||||||||||||||||||||
Qy   4734 GAAGGAGAGACCCCATCCTCACCTCATTATGGAACTTCTGGGGAGAAGGGCAGCAAAAAGGGACT 4793

Db   1927 GAAGGAGACCCCATCCTCA-CTCATTATGGAACTCTGGGGAGAGAGGGCAGCAAAAAGGGACT 1985
          ||||||||||||||||||| |||||||||||||| |||||||||| |||||||||||||||||
Qy   4794 CCTGCCACTCAGCAGCTGAAGTATCAACACCAGCCCTTGACCCTTGAGTCCTGGCTTTGG 4853

Db   1986 CCTGCCACTCAGCAGCTGAAGTATCAACACTAGCCCTTGACCCTTGAGTCCTGGCTTTGG 2045
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Qy   4854 CTGACCCTTCTTTGGGTCTCAGTTTCTTTCTTCTGCAAACAAGTTGCCATCTGGTTTTGCT 4913

Db   2046 CTGACCCTTCTTTGGGTCTCAGTTTTCTTTCTTCTGCAAACAAGTTGCCATCTGGTTTTGCT 2105
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy   4914 TCCAGCACTAAAGTAATTGGAACTTTGATGATGGCCTTTGCTGGGCCATTATGTGTCCATGCC 4973

Db   2106 TCCAGCACTAAAGTAAT--GGAACTTTGATGATGGCCTTTGCTGGGCCATTATGTGTCCATGCC 2165
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Qy   4974 AGGGATGCCACAGGGGGGCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATC 5033

Db   2166 AGGGATGCCACACGGGGGGCCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATC 2225
          ||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Qy   5034 TGGAGCTGGCAAGACCCTGCAGACCTCATAGAGCCTCATCTGGTTGGCCACAGCAGCCAA 5093

Db   2226 TGGAGCTGGCAAGACCCCTGCAGACCTCATAGAGCCTCATCTGGTTGGCCACAGCAGCACA 2285
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy   5094 GCCTAGAGACCCTCCGGATCCCATCCAGGCGGCAAAGAGGAAATAGGAGGGGACATGGAACCAT 5153

Db   2286 AGCCTAGAG-CCTCCGGATCCCATCCAGGCG-CAAAGAGGAA-TAGGAGGGGACATGGAACCAT 2345
          |||||||||  ||||||||||||||||||||  |||||||||  ||||||||||||||||||||
Qy   5154 TTGCCTCCTTGGGCTTGTCACAGGGGTGAGCCCCAAAAATTGGGGTTCAGCGTGGGAGGCCACG 5213

Db   2346 TTGCCTCT--GGGCTTGTCACAGGGGTGAGCCCCAAAAATTGGGGTTCAGCGTGGGAGGCCACG 2405
          ||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   5214 TGGATTCTTGGCTTTGTACAGAGAGATCTACAAGAGCAAGCCAACACAGAGTAAAGTGAAG 5273

Db   2406 TGGATTCTTGGCTTTGTACAGAGAGATCTACAAGAGCAAGCCAACACAGAGTAAAGTGGAAG 2465
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy   5274 GAAGTTTATTCAGAAAATAAAAGGAGTATCACACAGCTCTTTTAGAGAATTTGTCTACCAGGGCTT 5333

Db   2466 GAAGTTTATTCAGAAAAATAAAAGGAGTA--CACAGCTCTTTTAGAGAATTTGTCTACCAGGGCTT 2525
          |||||||||||||||| |||||||||||  |||||||||||||||||||||||||||||||||||
Qy   5334 TCCAGTTTTTACCAGAAAACCCCTATAAAATTAAAAAAATTTTTTACTTAAATTTAAGAAATTA 5393

Db   2526 TCCAGTTTTTACCAGAAAACCCCTATAAAATTAAAAAAATTTTTTACTTAAATTTAAGAAATTA 2585
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   5394 AAAAAAATACTAAAAAAAGAAAAAATGAAAAATAAAGGAATAAGGAGTTACCTAC 5444

Db   2586 AAAAAAATACAAAAAAAAGAAAAAATGAAAAATAAAGGAATAAGGAGTTACCTAC 2636
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
```

PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-DEC-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	09-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.

Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition -

XX
PS   Claim 1; SEQ ID NO 1745; 1297pp + Sequence Listing; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human reproductive system related antigens. These can be used
CC   in the prevention and treatment of reproductive system disorders,
CC   including cancer. The present sequence is a coding sequence of the
CC   invention.
XX
SQ   Sequence 578 BP; 130 A; 151 C; 166 G; 126 T; 5 other;

Query Match 8.9%; Score 484.4; DB 22; Length 578;
Best Local Similarity 99.4%; Pred. No. 5e-110;
Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy    1375 CAAAGACCTCACCCCAGAAGAGCCAAGATCCCCTGGTTCAAGAGAAGAGCTGGCCTGATTATGTG 1434
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 CAAAGACCTCACCCCAGAAGAGCCAAGATCCCCTGGTTCAAGAGAAGAGCTGGCCTGATTATGTG 61

Qy    1435 GAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCCTGGATACC 1494
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 GAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCCTGGATACC 121

Qy    1495 CAAGAAACCTCGCATAGTCATTCATACTGCAGGGGGGCCTGCTGGGAATTGGGAGTCAACACTGCC 1554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CAAGAAACCTCGCATAGTCATTCATACTGCAGGGGGGCTGCTGGGAATTGGGAAGTCAACACTGCC 181

Qy    1555 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGAGGGCCAGCTGTATGGGGACCGGCCTTCCAGCATGTC 1614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGAGGGCCAGCTGTATGGGGACCGGCTTCCAGCATGTC 241

Qy    1615 TTCTACTTCAGCTGCAGAGAGCTGGCCCAGTCCAAGTGGTCTCGCTAGCTCATC 1674
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 TTCTACTTCAGCTGCAGAGAGCTGGCCCAGTCCAAGTGGTCTCGCTGAGCTCATC 301

Qy    1675 GGAAAAAGATGGGACAGCCACTCCCGCCTCCCATTAGACAGATCCTGTCTAGGCCCAGAGCGG 1734
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 GGAAAAAGATGGGACAGCCACTCCCGGCTCCCATTAGACAGATCCTGTCTAGGCCCAGAGCGG 361

Qy    1735 CTGCTCTTCATCCTCGAATGGGTTGTAGATGAGGCCAGGATGGGTTCTTGCAGGAGGCCGAGTTCT 1794
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CTGCTCTTCATCCTCGAATGGTTGTAGATGAGGCCAGGATGGGTCTTGCAGGAGGCCGAGTTTCT 421

Qy    1795 GAGCTCTGTCTGCACTGGAGCCCACAG--CCGGCCGGATCGCACTGCTGCTGGGCAGTTTGCT 1853
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     422 GAGCTCTGTCTGCACTGGAGCCCACAGCCAGCCCGGGGGATGCACTGCTGCTGGGCANTTTGCT 481

Qy    1854 GGGGAAAACTATACTTCCC 1872
           |||||||||||||||||||
Db     482 GGGGAAAACTATACTTTCC 500
```

RESULT 13
ABL97037
ID   ABL97037 standard; cDNA; 578 BP.
XX
AC   ABL97037;
XX
DT   21-JUN-2002 (first entry)
XX
DE   Human testicular antigen encoding cDNA SEQ ID NO: 705.
XX
KW   Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW   reproductive system disorder; urinary system disorder; gene therapy;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disease; infection; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155317-A2.
XX
PD   02-AUG-2001.
XX

```
PF   17-JAN-2001;  2001WO-US01329.
XX
PR   31-JAN-2000;  2000US-0179065.
PR   04-FEB-2000;  2000US-0180628.
PR   24-FEB-2000;  2000US-0184664.
PR   02-MAR-2000;  2000US-0186350.
PR   02-MAR-2000;  2000US-0189874.
PR   16-MAR-2000;  2000US-0190076.
PR   17-MAR-2000;  2000US-0198123.
PR   18-APR-2000;  2000US-0205515.
PR   19-MAY-2000;  2000US-0209467.
PR   07-JUN-2000;  2000US-0214886.
PR   28-JUN-2000;  2000US-0215135.
PR   30-JUN-2000;  2000US-0216647.
PR   07-JUL-2000;  2000US-0216880.
PR   07-JUL-2000;  2000US-0217487.
PR   11-JUL-2000;  2000US-0217496.
PR   11-JUL-2000;  2000US-0218290.
PR   14-JUL-2000;  2000US-0220963.
PR   26-JUL-2000;  2000US-0220964.
PR   26-JUL-2000;  2000US-0225268.
PR   14-AUG-2000;  2000US-0225213.
PR   14-AUG-2000;  2000US-0225214.
PR   14-AUG-2000;  2000US-0225266.
PR   14-AUG-2000;  2000US-0225267.
PR   14-AUG-2000;  2000US-0225270.
PR   14-AUG-2000;  2000US-0225447.
PR   14-AUG-2000;  2000US-0225757.
PR   14-AUG-2000;  2000US-0225758.
PR   14-AUG-2000;  2000US-0225759.
PR   18-AUG-2000;  2000US-0226279.
PR   18-AUG-2000;  2000US-0226681.
PR   22-AUG-2000;  2000US-0226868.
PR   22-AUG-2000;  2000US-0227182.
PR   23-AUG-2000;  2000US-0227009.
PR   30-AUG-2000;  2000US-0228924.
PR   01-SEP-2000;  2000US-0229287.
PR   01-SEP-2000;  2000US-0229343.
PR   01-SEP-2000;  2000US-0229344.
PR   01-SEP-2000;  2000US-0229345.
PR   05-SEP-2000;  2000US-0229509.
PR   05-SEP-2000;  2000US-0229513.
PR   06-SEP-2000;  2000US-0230437.
PR   06-SEP-2000;  2000US-0230438.
PR   08-SEP-2000;  2000US-0231242.
PR   08-SEP-2000;  2000US-0231243.
PR   08-SEP-2000;  2000US-0231244.
PR   08-SEP-2000;  2000US-0231413.
PR   08-SEP-2000;  2000US-0231414.
PR   08-SEP-2000;  2000US-0232080.
PR   08-SEP-2000;  2000US-0232081.
PR   12-SEP-2000;  2000US-0231968.
PR   14-SEP-2000;  2000US-0232397.
PR   14-SEP-2000;  2000US-0232398.
PR   14-SEP-2000;  2000US-0232399.
PR   14-SEP-2000;  2000US-0232400.
PR   14-SEP-2000;  2000US-0232401.
PR   14-SEP-2000;  2000US-0233063.
PR   14-SEP-2000;  2000US-0233064.
PR   14-SEP-2000;  2000US-0233065.
PR   21-SEP-2000;  2000US-0234223.
PR   21-SEP-2000;  2000US-0234274.
PR   25-SEP-2000;  2000US-0234997.
PR   25-SEP-2000;  2000US-0234998.
PR   26-SEP-2000;  2000US-0235484.
PR   27-SEP-2000;  2000US-0235834.
PR   27-SEP-2000;  2000US-0235836.
PR   29-SEP-2000;  2000US-0236327.
PR   29-SEP-2000;  2000US-0236367.
PR   29-SEP-2000;  2000US-0236368.
PR   29-SEP-2000;  2000US-0236369.
PR   29-SEP-2000;  2000US-0236370.
PR   02-OCT-2000;  2000US-0236802.
PR   02-OCT-2000;  2000US-0237037.
PR   02-OCT-2000;  2000US-0237038.
PR   02-OCT-2000;  2000US-0237039.
PR   02-OCT-2000;  2000US-0237040.
PR   13-OCT-2000;  2000US-0239935.
PR   20-OCT-2000;  2000US-0239937.
PR   20-OCT-2000;  2000US-0240960.
PR   20-OCT-2000;  2000US-0241221.
PR   20-OCT-2000;  2000US-0241785.
PR   20-OCT-2000;  2000US-0241786.
PR   20-OCT-2000;  2000US-0241787.
PR   20-OCT-2000;  2000US-0241808.
PR   20-OCT-2000;  2000US-0241809.
PR   20-OCT-2000;  2000US-0241826.
PR   01-NOV-2000;  2000US-0244617.
PR   08-NOV-2000;  2000US-0246474.
PR   08-NOV-2000;  2000US-0246475.
PR   08-NOV-2000;  2000US-0246476.
PR   08-NOV-2000;  2000US-0246477.
PR   08-NOV-2000;  2000US-0246478.
PR   08-NOV-2000;  2000US-0246523.
PR   08-NOV-2000;  2000US-0246524.
PR   08-NOV-2000;  2000US-0246525.
PR   08-NOV-2000;  2000US-0246526.
PR   08-NOV-2000;  2000US-0246527.
PR   08-NOV-2000;  2000US-0246528.
PR   08-NOV-2000;  2000US-0246532.
PR   08-NOV-2000;  2000US-0246609.
PR   08-NOV-2000;  2000US-0246610.
PR   08-NOV-2000;  2000US-0246611.
PR   08-NOV-2000;  2000US-0246613.
PR   17-NOV-2000;  2000US-0249207.
PR   17-NOV-2000;  2000US-0249208.
PR   17-NOV-2000;  2000US-0249209.
PR   17-NOV-2000;  2000US-0249210.
PR   17-NOV-2000;  2000US-0249211.
PR   17-NOV-2000;  2000US-0249212.
PR   17-NOV-2000;  2000US-0249213.
PR   17-NOV-2000;  2000US-0249214.
PR   17-NOV-2000;  2000US-0249215.
PR   17-NOV-2000;  2000US-0249216.
PR   17-NOV-2000;  2000US-0249217.
PR   17-NOV-2000;  2000US-0249218.
PR   17-NOV-2000;  2000US-0249244.
PR   17-NOV-2000;  2000US-0249245.
PR   17-NOV-2000;  2000US-0249264.
PR   17-NOV-2000;  2000US-0249265.
PR   17-NOV-2000;  2000US-0249297.
PR   17-NOV-2000;  2000US-0249299.
PR   17-NOV-2000;  2000US-0249300.
PR   01-DEC-2000;  2000US-0250160.
PR   01-DEC-2000;  2000US-0250391.
PR   05-DEC-2000;  2000US-0251030.
PR   05-DEC-2000;  2000US-0251479.
PR   06-DEC-2000;  2000US-0251719.
PR   06-DEC-2000;  2000US-0251856.
PR   05-DEC-2000;  2000US-0251868.
PR   08-DEC-2000;  2000US-0251869.
PR   08-DEC-2000;  2000US-0251989.
PR   08-DEC-2000;  2000US-0251990.
PR   11-DEC-2000;  2000US-0254097.
PR   05-JAN-2001;  2001US-0259678.
XX
PT   Nucleic acids encoding 973 human testicular antigen polypeptides,
XX
DR   WPI; 2001-483232/52.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
```

PT   useful for preventing, diagnosing and/or treating testicular cancer -
XX
PS   Claim 1; SEQ ID NO 705; 766pp; English.
XX
CC   The present invention provides the protein and coding sequences of 973
CC   human testicular antigens, and fragments of their genomic sequences. The
CC   sequences can be used in the treatment of cardiovascular, urinary system,
CC   reproductive system, immune, respiratory, neurological and
CC   gastrointestinal disorders, infections, and particularly cancer,
CC   especially testicular cancers. The present sequence is a cDNA of the
CC   invention.
XX
SQ   Sequence 578 BP; 130 A; 151 C; 166 G; 126 T; 5 other;

Query Match   99.4%; Score 484.4; DB 23; Length 578;
Best Local Similarity   99.4%;
Matches 496; Conservative   0; Mismatches   2; Indels   1; Gaps   1;
                                   Pred. No.5e-110;

Qy   1375 CAAAGACCTTCACCCCAGAAGCCAAGCAGATCCCCTGGTTCAAGAGGAAGCTGGGCCTGATTATGTG 1434
           |||||||||| |||||||||||||||||| |||||||||||||| ||||||||||||||||||||||
Db   2 CAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTTCAAGAGAAGCTGGGCCTGATTATGTG 61

Qy   1435 GAGGAGAATGGAGGACATTTAATTGAGATCAGAGACTTATTTGGGCCCAGGCCTGGATACC 1494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GAGGAGAATGGAGGACATTTAATTGAGATCAGAGACTTATTTGGGCCCAGGCCTGGATACC 121

Qy   1495 CAAGAAACCTCGCATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAAGTCAAACACTGGCC 1554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   122 CAAGAAACCTCGCATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAAGTCAACACTGGCC 181

Qy   1555 AGGGCAGGTGAAGGAAGCCTTGGGGAGAGAGGGCCAGCGTTCCAGCAATGTC 1614
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   182 AGGGCAGGTGAAGGAAGCCTGGGGGAGAGAGGGCCAGCGTTCCAGCAATGTC 241

Qy   1615 TTCTACTTCAGCTGCAGAGAGACTGGCCCAGTCCAAGGTGGTGAGTCTCGCTGAGCTCATC 1674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 TTCTACTTCAGCTGCAGAGAGACTGGCCCAGTCCAAGGTGGTGAGTCTCGCTGAGCTCATC 301

Qy   1675 GGAAAAAGATGGGACCAGCCACTCCGGCCTCCCATTAGACAGATCCTGTCTAGGCCACCAGAGCGG 1734
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   302 GGAAAAAGATGGGACCAGCCACTCCGGCCTCCCATTAGACAGATCCTGTCTAGGCCACCAGAGCGG 361

Qy   1735 CTGCTCTTCATCCTCGATGGTGTAGATGAGCCACGAGATGGGTCTTCCAGGAGGCCGAGTTCT 1794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 CTGCTCTTCATCCTCGATGGTGTAGATGAGCCACGAGATGGGTCTTCCAGGAGGCCGAGTTCT 421

Qy   1795 GAGCTCTGTCTGCACTGGGAGCCAGCCACAG-CCGGCGGGATGCACTGCTGGGCCAGTTTGCT 1853
           |||||||||||||||||| ||||||||| ||| ||| |||||||||||||||||||||||||||
Db   422 GAGCTCTGTCTGCACTGGGAGCCAGCCCCAGGCCCGGGGGGATGCACTGCTGGGCCANTTTGCT 481

Qy   1854 GGGGAAAAACTATACTTCCC 1872
           |||||||||||||||||| ||
Db   482 GGGGAAAAACTATACTTTCC 500

RESULT 14

ABV97890/c
ID   ABV97890 standard; cDNA; 487 BP.
XX
AC   ABV97890;
XX
DT   14-JAN-2003 (first entry)
XX
DE   Human pancreatic cancer expressed cDNA SEQ ID NO 3298.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.

XX
PF   30-JAN-2002; 2002WO-US02781.
XX
PR   30-JAN-2001; 2001US-265305P.
PR   31-JAN-2001; 2001US-265682P.
PR   09-FEB-2001; 2001US-267568P.
PR   21-MAR-2001; 2001US-278651P.
PR   28-APR-2001; 2001US-287112P.
PR   16-MAY-2001; 2001US-291631P.
PR   12-JUL-2001; 2001US-305494P.
PR   20-AUG-2001; 2001US-313999P.
PR   27-NOV-2001; 2001US-333626P.
XX
DR   WPI; 2002-627435/67.
XX
PT   New isolated polynucleotide and pancreatic tumor polypeptides, useful
PT   for diagnosing, preventing and/or treating cancer, particularly
PT   pancreatic cancer -
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
PS   Claim 1; SEQ ID NO 3298; 300pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising: (a)
CC   any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
CC   (b) complements of (a); (c) sequences consisting of at least 20
CC   contiguous residues of (a); (d) sequences that hybridize to (a), under
CC   moderately stringent conditions; (e) sequences having at least 75% or 90%
CC   identity to (a); or (f) degenerate variants of (a). Polypeptides
CC   (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
CC   detect cancer in a patient comprising polypeptides,
CC   polynucleotides, antibodies, fusion proteins, T cell populations and
CC   antigen presenting cells expressing the polypeptide are useful in
CC   treating pancreatic cancer and stimulating an immune response. The
CC   polynucleotides can be used as probes or primers for nucleic acid
CC   hybridisation, in the design and preparation of ribozyme molecules for
CC   inhibiting expression of the tumour polypeptides and proteins in the
CC   tumour cells, in vaccines and for gene therapy.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 487 BP; 114 A; 130 C; 133 G; 109 T; 1 other;

Query Match   99.4%; Score 482.8; DB 24; Length 487;
Best Local Similarity   99.4%;
Matches 484; Conservative   0; Mismatches   3; Indels   0; Gaps   0;
                                   Pred. No.1.1e-109;

Qy   4744 CATCCTCCACCTCATTATGGAACTCTGGGAGAGGGCAGCAAAAAGGGACTCTGCCACTC 4803
           |||||||| ||| ||||||| ||||||||||||| ||||||||||||||||||||||||||
Db   487 CATCCTCACCTTATTATGNAACTCTGGGAGAAGGGCAGCAAAAAGGGACTCTGCCACTC 428

Qy   4804 AGCAGCTGAAGTATCAACACACCAGCCCTTGACCCTTGAGTCCTGGCTTTGGCTGACCCTTC 4863
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   427 AGCAGCTGAAGTATCAACAC-CAGCCCTTGACCCTTGAGTCCTGGCTTTGGCTGACCCTTC 368

Qy   4864 TTTGGGTCTCAGTTTCTTCTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTA 4923
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 TTTGGGTCTCAGTTTCTTCTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTA 308

Qy   4924 AAGTAATGGAACTTTGATGATGTCCTTTGCTGGGCATTATGTGTCCAATGCCAGGGATCCA 4983
           ||||||||||||||||||||||| |||||||||||||||||||||||| |||||||||||||
Db   307 AAGTAATGGAACTTTGATGATGCCTTTGCTGGGCATTATGTGTCCATGCCAGGGATCCA 248

Qy   4984 CAGGGGGCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTTCATTCTGGAGCTGGC 5043
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   247 CAGGGGGCCCCAGTCCAGGTGGCATAACAGCATCTCAGGGAATGTTCATTCTGGAGCTGGC 188

Qy   5044 AAGACCCCTGCAGACCTCATAGAGCTTCATCTGGTGGCCACAGCAGCCAAGCCTAGAGCC 5103
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   187 AAGACCCCTGCAGACCTTCATAGAGCTTCATCTGGTGGCCACAGCAGCCAAGCCTAGAGCC 128

```
Qy  5104 CTCCGGATCCATCCAGGCGGCCAAAGAGAGGAGGAATAGGAGGGACATGGGAACCATTTGCCTCTCGG 5163
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTCCGGATCCATCCAGGCGGCCAAAGAGAGGAGGAATAGGAGGGACATGGGAACCATTTGCCTCTCGG   68

Qy  5164 CTGTGTCACAGGGTGAGCGCCCAAAATTGGGGTTCAGCGTGGGAGGAGGCCACGTGGGATTCTTG 5223
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 CTGTGTCACAGGGTGAGCGCCCAAAATTGGGGTTCAGCGTGGGAGGAGGCCACGTGGGATTCTTG    8

Qy  5224 GCTTTGT 5230
         |||||||
Db     7 GCTTTGT 1
```

RESULT 15

```
ID   ABV97891 standard; cDNA; 416 BP.
XX
AC   ABV97891;
XX
DT   14-JAN-2003  (first entry)
XX
DE   Human pancreatic cancer expressed cDNA SEQ ID NO 3299.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour; gene; ss.
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US02781.
XX
PR   30-JAN-2001; 2001US-265305P.
PR   31-JAN-2001; 2001US-265682P.
PR   09-FEB-2001; 2001US-267568P.
PR   21-MAR-2001; 2001US-278651P.
PR   28-APR-2001; 2001US-287112P.
PR   16-MAY-2001; 2001US-291631P.
PR   12-JUL-2001; 2001US-305484P.
PR   20-AUG-2001; 2001US-313999P.
PR   27-NOV-2001; 2001US-333626P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR   WPI; 2002-627435/67.
XX
PT   New isolated polynucleotide and pancreatic tumor polypeptides, useful
PT   for diagnosing, preventing and/or treating cancer, particularly
PT   pancreatic cancer -
XX
PS   Claim 1; SEQ ID NO 3299; 300pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising: (a)
CC   any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
CC   (b) complements of (a); (c) sequences consisting of at least 20
CC   contiguous residues of (a); (d) sequences that hybridize to (a), under
CC   moderately stringent conditions; (e) sequences having at least 75% or 90%
CC   identity to (a); or (f) degenerate variants of (a). Polypeptides
CC   (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
CC   detect cancer in a patient and compositions comprising polypeptides,
CC   polynucleotides, antibodies, fusion proteins, T cell populations and
CC   antigen presenting cells expressing the polypeptide are useful in
CC   treating pancreatic cancer and stimulating an immune response. The
CC   polynucleotides can be used as probes or primers for nucleic acid
CC   hybridisation, in the design and preparation of ribozyme molecules for
CC   inhibiting expression of the tumour polypeptides and proteins in the
CC   tumour cells, in vaccines and for gene therapy.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 416 BP; 101 A; 113 C; 114 G; 88 T; 0 other;

Query Match             99.8%;  Score 414.4;  DB 24;  Length 416;
Best Local Similarity   99.8%;  Pred. No. 1.1e-92;
Matches  415;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy  4815 TATCAACACCAGCCCTTGACCCTTGAGTCCTGGGCTTTGGCGTGACCCTTCTTTGGGGTCTCA 4874
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   416 TATCAACACCAGCCCTTGACCCTTGAGTCCTGGGCTTTGGCGTGACCCTTCTTTGGGGTCTCA  357

Qy  4875 GTTTCTTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAA 4934
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 GTTTCTTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAA  297

Qy  4935 CTTTGATGATGCCTTTGCTGGGCATTATGTGTCCATGCCAGGGATGTCCACAGGGGGGCCC 4994
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   296 CTTTGATGATGCCTTTGCTGGGCATTATGTGTCCATGCCAGGGATGTCCACAGGGGGGCCC  237

Qy  4995 AGTCCAGGTGGCCTRAACAGCATCTCAGGGGAATGTCCATCTGGGAGTGGGCAAGACCCCTGC 5054
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 AGTCCAGGTGGCATRAACAGCATCTCAGGGGAATGTCCATCTGGGAGTGGGCAAGACCCCTGC  177

Qy  5055 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCTAGAGCCCTCCGGATCCC 5114
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCTAGAGCCCTCCGGATCCC  117

Qy  5115 ATCCAGGCGCAAAGAGGAATAGGAGGGACACATGGGAACCATTTGCCTCTCGGCGTGGTCACAG 5174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116 ATCCAGGCGCAAAGAGGAATAGGAGGGACACATGGGAACCATTTGCCTCTCGGCGTGGTCACAG   57

Qy  5175 GGTGAGCCCCAAAATTGGGGTTCAGCGTGGGAGGAGGCCACGTGGGATTCTTGGCTTTGT 5230
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    56 GGTGAGCCCCAAAATTGGGGTTCAGCGTGGGAGGAGGCCACGTGGGATTCTTGGCTTTGT    1

Search completed: January 29, 2004, 03:29:21
Job time : 1357.51 secs
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 29, 2004, 03:01:52 ; Search time 10849.3 seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcgcctggagaggt..........aggaataagaagttaacctac 5444

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

```
 1: em_estba:*       16: em_estom:*
 2: em_esthum:*      17: em_hum:*
 3: em_estim:*       18: em_gss_inv:*
 4: em_estmu:*       19: em_gss_pln:*
 5: em_estov:*       20: em_gss_vrt:*
 6: em_estpl:*       21: em_gss_fun:*
 7: em_estro:*       22: em_gss_mam:*
 8: em_htc:*         23: em_gss_mus:*
 9: em_est1:*        24: em_gss_pro:*
10: gb_est2:*        25: em_gss_rod:*
11: gb_htc:*         26: em_gss_phg:*
12: gb_est3:*        27: em_gss_vrl:*
13: gb_est4:*        28: gb_gss1:*
14: gb_est5:*        29: gb_gss2:*
15: em_estun:*
```

SUMMARIES

No.	Score	% Match	Query Length	DB ID	ID	Description
1	994	18.3	1201	9	AL541851	AL541851 AL541851
2	983.2	18.5	1201	13	BX341134	BX341134 BX341134
3	898.4	16.5	1201	13	BX341135	BX341135 BX341135
4	819.8	15.1	1055	13	BQ072636	BQ072636 AGENCOURT
5	814.4	15.0	1046	12	BM549709	AGENCOURT
6	787.4	14.5	1110	13	BX457482	BX457482
7	767	14.1	768	12	BI518887	603061807
8	749.4	13.8	894	14	CD515485	AGENCOURT
9	727	13.4	791	10	BG682080	602630057
10	722.4	13.3	764	13	BQ576345	UI-H-BI1-
11	671	12.3	686	14	CA308213	UI-H-FT1-
12	633.6	11.6	672	12	BI911134	603067546
13	633	11.6	728	12	BI908442	603067392
14	628.4	11.5	655	12	BQ003173	UI-H-EI1-
15	623.6	11.5	655	12	BM152462	TCBAP1E90
16	615	11.3	615	12	BM791087	K-EST0071
17	575	10.6	592	13	BU639947	UI-H-FL1-
18	566	10.4	566	10	BF059271	7k57c12.x
19	558.4	10.3	560	12	BM152839	TCBAP2Q12
20	544.4	10.0	554	12	BM193542	TCBAP2Q14
21	520.8	9.6	524	9	AW173055	xj82f03.x
22	520.4	9.6	523	12	BM153182	TCBAP1Q14
23	509	9.3	595	13	BQ355077	MR2-HT116
24	502.6	9.2	514	9	AV682872	AV682872
25	495.4	9.1	580	12	BM152807	TCBAP1Q12
26	494	9.1	494	9	AW593727	hg21g12.x
27	489.6	9.0	806	12	BI519523	603061807
28	485	8.9	494	12	BM151763	TCBAP1E10
29	483.4	8.9	500	10	BE775300	QV0-UM009
30	476.4	8.8	633	12	BI909022	603070167
31	475	8.7	509	12	BG774227	60262231
32	474.6	8.7	521	9	AI819334	wg55h07.x
33	469.8	8.6	504	9	AI742596	wg56h07.x
34	469	8.6	469	9	AI825353	wb17d08.x
35	458	8.4	458	10	BE464469	h2l7e05.x
36	456.8	8.4	511	13	BX474531	DKFZp686G
37	449	8.2	449	9	AI365110	qx96b04.x
38	445.4	8.2	461	9	AW291804	UI-H-BI2-
39	445.2	8.2	487	12	BI028123	CM4-MT028
40	441	8.1	464	9	AW291769	UI-H-BI2-
41	438	8.0	441	9	AI339112	qt06b07.x
42	434.4	8.0	466	12	BM193358	TCBAP1E96
43	434.2	8.0	791	28	BZ129303	CH230-345
44	424	7.8	425	12	AI439770	ti61c01.x
45	423.2	7.8	489	12	BM151640	TCBAP1E10

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
AL541851

```
LOCUS       AL541851   1201 bp   mRNA   linear   EST 12-MAY-2003
DEFINITION  AL541851 Homo sapiens PLACENTA Homo sapiens cDNA clone CSODE007YD22
            5-PRIME, mRNA sequence.
ACCESSION   AL541851
VERSION     AL541851.2  GI:30546421
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B.; Gruber,C.; Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12873320.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8396. For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CSODE007DB11QP1&cluster=8396.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL ;
```

http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
Faraday Avenue Genoscope Sequence ID : CS0DE007DB11QP1.

FEATURES
source   1..1201
                Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DE007YD22"
                /tissue_type="PLACENTA"
                /clone_lib="Homo sapiens PLACENTA"
                /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoR V sites of the pCMVSPORT 6 vector.
                Library was not normalized."

BASE COUNT   264 a   370 c   301 g   225 t   41 others
ORIGIN

Query Match
Best Local Similarity   96.2%;   Score 994;   DB 9;   Length 1201;
Matches 1008;   Conservative 17;   Mismatches 22;   Indels 1;   Gaps 1;

Qy   5 CAGGGCCTGGAGAGGTCTGAAGAAAACCTGGGAGGCCAGCAGCGAGGCCCGGGGGCTTCCACTCTGGGT 64
            | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   32 CCGGTCCGGGAATTCCCGGGATGGAGAAACCTGGGAGGCCAGCAGCGAGGCCCGGGGGCTTCCACTCTGGGT 91

Qy   65 TCTTGAAAGCCCATTCCCTGCTCTGCGGGCTCCTCCCCCACCTCTTCTTCTTCAGCCTTGCAG 124
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   92 TCTTGAAAGCCCATTCCCTGCTCTGCGGGCTCCTCCCCCACCTCTTCTTCTTCAGCCTTGCAG 151

Qy   125 CTCAAGGGGTTGATCTCAGGAGGTCCAGGAGGAAGAAATCTGAGGAACACAGA 184
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 CTCAAGGGGTTGATCTCAGGAGGTCCAGGACCCAGGAGAGGAGAGAGAATCTGAGGAACACAGA 211

Qy   185 ACAGTGAGGGTTGCCCACACCCCAATCTCCCGTCACCAACATCTCCCCTACCCTACCCTC 244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212 ACAGTGAGGGTTGCCCACACCCCAATCTCCCGTCACCAACATCTCCCCTACCCTACCCTC 271

Qy   245 CCTGCCCTGGCCCCTGGACCCCATCCCATACTCAGCTGACTTCTTCCAGTGTCT 304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   272 CCTGCCCTGGCCCCTGGACCCCATCCCATACTCAGCTGACTTCTTCCAGTGTCTT 331

Qy   305 TGCAGGCCCCTCTGGGCTCCTCCCCCTGGCTTTCCTACCACTCCCCCTCTATCGGC 364
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   332 TGCAGGCCCCTCTGGGCTCCTCCCCCTGGCTTTCCTACCACTCCCCCTCTATCGGC 391

Qy   365 GTCTATCTGTAGGTGTGCCCTGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGAGACG 424
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   392 GTCTATYTGTAGGTGTGCCCTGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGAGACG 451

Qy   425 GTAAGAGCCAAGGCAAAGGACACACTGTTTCTTCGTCCTGCCTGCTATCCCTACCACCACTTGG 484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 GTAAGAGCCAAGGACAGCACTGTTCTTCGTCCTGCCTGATACCCTACCACCACTTGG 511

Qy   485 GAACATCCCCACGACACCCTCTTAAACTCCGGACAGGATGGCTGGGGGGGAGGCCTGGGGGCC 544
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   512 GAACATCCCCAGACACCCTCTTAAACTCCGGAGCAGAGAGATGGCTGGGCGAG-CTGGGGCC 570

Qy   545 GCCTGGGCCTGTTACTTGGAGTTCTTGAAGAAGAGGAGCTGAAGGAGTTCCAGCTTCTGC 604
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   571 GCCTGGCCTGTTACTTGGAGTTCTTGAAGAAGAGGAGCTGAAGGAGTTCCAGCTTCTGC 630

Qy   605 TCCCCAATAAAAGCCCACTCCCAGGAGAGGTCTTCAGGGGTGAGACACCCGCTCAGCCGAGAGGAGA 664
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   571 TCCCCAATAAAAGCCCACTCCCAGGAGACTCTTCAGGGGTGAGACACCCGCTCAGCCGAGAGGAGA 664

Qy   631 TCGGCCAATAAAGCCACTCCCAGGAGACTCTTTCCGGTGAGACACCCGGCTCAGCCGAGAGGA 690
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   665 CGAGTGGGCATGGAGGTGGCCTCGTACCTGGGG 750

Qy   665 CGAGTGGGCATGGAGGTGGCCTCGTACCTGGTGGTGGCTCAGTATGGGAGCAGCGCGGCCGCCTGG 724
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   691 CGAGTGGGCATGGAGGTGGCCTCGTACCTGGGG 750

Qy   725 ACCTAGCCCTCCATACCTGGGAGGCAGATGGGAGGCTGAGGTCACTGGGCCCAAGCCCAGG 784
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   751 ACCTAGCCCTCCATACCTGGGAGGCAGATGGGAGGCTGAGGTCACTGGGCCCAAGCCCAGG 810

RESULT 2
BX341134/c

LOCUS   BX341134   1201 bp   mRNA   linear   EST   01-MAY-2003
DEFINITION   BX341134 Homo sapiens cDNA clone CS0DJ012YC02 3-PRIME, mRNA sequence.
               Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION   BX341134
VERSION   BX341134.1   GI:30309318
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 1201)
AUTHORS   Li, W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. This sequence belongs to sequence cluster 8396.r For
               more information about this cluster, see
               http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ012BB01NP1
               &cluster=8396.r. Contact : Feng Liang Email : fliang@lifetech.com
               URL : http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
               Faraday Avenue Genoscope sequence ID : CS0DJ012BB01NP1.

FEATURES
source   1..1201
                Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DJ012YC02"
                /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                /cell_line="JURKAT"
                /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                10-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and cloned into the Not I and EcoR V
                sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT   263 a   286 c   288 g   313 t   51 others
ORIGIN

Query Match
Best Local Similarity   95.6%;   Score 983.2;   DB 13;   Length 1201;
Matches 1000;   Conservative 9;   Mismatches 37;   Indels 0;   Gaps 0;

Qy   785 AAGGGGCAGGGCCACTCTCCTTCCTCATTCCCTACAGCCCACCCAAGTGAACCCCACCTGGGGGTCTC 844
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   811 AAGGGGCAGGGCCACTCTCCCTCATTCCCTACAGCCCCAAGTGAACCCCACCTGGGGGTCTC 870

Qy   845 CCAGCCAACCCACCTCCACCGCAGTGTCCTGATCCATGAATTGAATTGCCCGGCCGCC 904
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   871 CCAGCCAACCCACCTCCACCGCAGTGTCTAATGCCCCTACCACCCCAGGTGAACCCCAC 930

Qy   905 GCACCCCAGGGGCTCAGGAGGAAGGGGTTTTGAGGACAGCTGGCTGACACATCTGGACGCGCCGCT 964
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   931 GCACCCAGGGGCTCAGAGAGRAAAGGGTTTTGAGAMARCTGCCTGACACATCTGGACGCGCCGGT 990

Qy   965 GGAGGAGAAATCTTGCCTCACTCCTCTACCAAGGTCTTCCAAGTCCCCAGACCCATGAGT 1024
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   991 GGRGAGAAATCTTYKCCTYACTYCTYCTACCAAGGTCTCTTYCAAGGTCCCCARACCCATGAGT 1050

Qy   1025 CTTCCAAGCCAGAGGAGTCACCCAAGCCCC 1052
            ||||||| |:| |||||||||| |||
Db   1051 TTCCAASCAGGRGRKTACCCAMGCCCCC 1078

FEATURES
source   1..1201
                Location/Qualifiers

Query Match
Best Local Similarity   18.1%;   Score 983.2;   DB 13;   Length 1201;
Matches 1000;   Conservative 9;   Mismatches 37;   Indels 0;   Gaps 0;

Qy   4366 GGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAAGAACAAGAAAAGATGAGACTCTGGTG 4425
            ||:|| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1057 GTCVACKTGCCACTTGGATCRGGATCCAGCTGCAAGTRAARACAAGAAAGWGAGACTCTGG 998

```
DEFINITION   BX341135 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
             Homo sapiens cDNA clone CS0DJ012YC02 5-PRIME, mRNA sequence.
ACCESSION    BX341135
VERSION      BX341135.1  GI:30310958
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1201)
  AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE      Full-length cDNA libraries and normalization
  JOURNAL    Unpublished
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             This sequence was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 8396.r For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS0DJ012BB01QP1&cluster=8396.r. Contact :
             Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0DJ012BB01QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ012YC02"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT        297 a        329 c        312 g        217 t       46 others
ORIGIN
```

Query Match   16.5%;   Score 898.4;   DB 13;   Length 1201;
Best Local Similarity   97.4%;
Matches 937; Conservative   7;   Mismatches   13;   Indels   5;   Gaps   3;
                                  Pred. No. 6.2e-111;

```
Qy   4426 TGGGAGGCCTTGGTGAAACCAGGAGATCTCATGCCTGCGAACTACTCACTCTCCAGCC 4485
Db    997 TGTGGAGGSCCTTGTGAAACCARGAGATCTCATGCCTGCGAACTACTCACTCTCCAGCC 997

Qy   4486 CGCATAGCCGTACCTTCACCTTCTGGATGCCCGCCCGCAGTTGCTGCACTTTGTTGGACCAGTAT 4545
Db    937 CGCATAGCCGTACCTTCACCTTCTGGATGCCCGCCCGCAGTTGCTGCACTTTGTTGGACCAGTAT 938

Qy   4546 CGAGAGCAGCTGATAGCCCGAGTGACATGCGTGGAGGTTGTTCTTTGGACAAACTGACCAGTGA 4605
Db    877 CGAGAGCAGCTGATAGCCCGAGTGACATGCGTGGAGGTTGTTCTTTGGACAAACTGATGGA 878

Qy   4606 CAGGTGCTGAGCCAGGAGCAGGTACGCAGTAGCAGTAGAGGGTGCTGGCTTGAGAACAGGCCCAGCCAG 4665
Db    817 CAGGTGCTGAGCCAGGAGCAGGTACGCAGTAGCAGTAGAGGGTGCTGGCTTGAGAACAGGCCCAGCCAG 758

Qy   4666 ATGCGGAAGCTGTTGTTCAGCTTGAGCCAGTCTCTGGGACCCGGAAGTGAAAAGATGGGACTCTAC 4725
Db    757 ATGCGGAAGCTGTTGTTCAGCTTGAGCCAGTCTCTGGGACCCGGAAGTGAAAAGATGGGACTCTAC 698

Qy   4726 CAAGCCCTGAAGGAGACGACCCATCCTCACCTCATTATGGAACTCTGGAGGAAGGGCAGCAGCCAAA 4785
Db    697 CAAGCCCTGAAGGAGGACCCCATCCTATTATGGGAGAACTCTGGGAGAAGGGCAGCAAA 638

Qy   4786 AAGGGACTCCTGCCACTCAGCAGCTGAAGTATCAACACAGTCCCTTGACCCCTTGAGTGTCCT 4845
Db    637 AAGGGACTCCTGCCACTCAGCGAGCTGAAGTATCAACACAGTCCCTTGACCCCTTGAGTGTCCT 578

Qy   4846 GGCTTGGCTGACCCTTCTTTGGGCGTCTCAGTTTCTTCTCTCGCAAACAAGTTGCCATCTG 4905
Db    577 GGCTTGGCTGACCCTTCTTTGGGCGTCTCAGTTTCTTCTCTCGCAAACAAGTTGCCATCTG 518

Qy   4906 GTTTGCCTTCCAGCACTCACTAAAGTAATGGAACTTTGAATGGATGCCCTTTGCTTGGTCATTATGTG 4965
Db    517 GTTTGCCTTCCAGCACTAAAGTAATGGAACTTTGATGATGCCCTTTGCTCTGGCCATTATGTG 458

Qy   4966 TCCATGCCCAGGGATGCCACAGGGGGCGCCCCAGTCAGTGGCTAACATGCATCTCAGGGAA 5025
Db    457 TCCATGCCCAGGGATGCCACAGGGGGCGCCCCAGTCAGTGGCCTAACAGCATCTCAGGGAA 398

Qy   5026 TGTCCATCTCGAGCTGGCCAAGACCCCTGCAGACCTCATAGAGCCTTCATCTGGTGGCCACA 5085
Db    397 TGTCCATCTCGAGCTGGCCAAGACCCCTGCAGACCTCATAGAGCCTTCATCTGGTGGCCACA 338

Qy   5086 GCAGCCAAGCCTAGAGCCCTCCCGGATCCCATCCAGCGCGCAAAAGAGGGAATAGGAGGGACAT 5145
Db    337 GCAGCCAAGCCTAGAGCCCTCCCGGATCCCATCCAGCGCGCAAAAGAGGGAATAGGAGGGACAT 278

Qy   5146 GGAACCATTGGCCTCTGGCTGTGTCACAGGGTGAGCCCCAAAATTGGGGTTCAGCGTGGG 5205
Db    277 GGAACCATTGGCCTCTGGCTGTGTCACAGGGTGAGCCCCAAAATTGGGGTTCAGCGTGGG 218

Qy   5206 AGGCCACGTGGATTCTGGATTCTTGGGCTTTGTGTACAGAGAGACTACAGAGCCAACACAGAGTAA 5265
Db    217 AGGCCACGTGGATTCTGGATTCTTGGGCTTTGTGTACAGAGAGACTACAGAGCCAACACAGAGTAA 158

Qy   5266 AGTTGGAAAGGAAGTTTTATTCTGGAAAAAATAAAGAGAGTAATCACAGCTCTTTAGAATTTGTCTA 5325
Db    157 AGTTGGAAAGGAAGTTTATTCAGAAAAAATAAAGAGAGTATCACAGCTCTTTAGAATTTGTCTA 98

Qy   5326 GCAGGCTTTCCAGTTTTTACCAGGAAAAACCCCTATAAATTAAAAAATTTTTTTTACTTAAAATTT 5385
Db     97 GCAGGCTTTCCAGKTTTTACCAGGAAAAACCCCTATAAATTAAAAAATTTTTTTTACTTAAAATTT 38

Qy   5386 AAGAATTAAAAAAAATACAAAAAAGAA 5411
Db     37 AAGAATTAAAAAAAATACAAAAAAGAR 12
```

```
Qy    773 CCCAAGCCCAGGAAGGGGCAGGCCACTCACTCCTCCTTCATTCCCTCAGGCGCCAGTGAACCCC 832
Db     57 CCCGGGATCAGGAAGGGGCAGGCCACTCTCCCCTCATTCCCTCAGGCCACAGTGAACCCC 116

Qy    833 ACCTGGGGTCTCCCACCACACCTCCACCGCAGTGCTAATGCCCTGGATCCTGGATCCATGAAT 892
Db    117 ACCTGGGGTCTCCCACCACACCTCCACCGCAGTGCTAATGCCCTGGATCCTGGATCCATGAAT 176

Qy    893 TGCCGCGGGGGTGCACCCAGGCCTCAGAGAAGAAGGGTTTTTGAGACACAGCTGGTTGG 952
Db    177 TGCCGCGGGGGTGCACCCAGGCCTCAGAGAAGAAGGGTTTTTGAGACACAGCTGACACAT 236

Qy    953 CTGGACGCGCGTGGAGAGAAATCTCGCCTCACTCCTTCCTACCAAGCTCTTCCAAGCTCCC 1012
Db    237 CTGGACGCGCGTGGAGAGAAATCTCGCCTCACTCCTTCCTACCAAGCTCTTCCAAGCTCCC 296

Qy   1013 CAGACCATGAGTCTCCAAGCCAGGACTCACCCAAGCACCAATCCACACAGAGCTCC 1072
Db    297 CAGACCATGAGTCTCCAAGCCAGGAGTCACCCAGCCACCAAGCGCCCCACATCCACAGAGCTCCTGGGA 356

Qy   1073 GGAGCTGGGGGGATCCCCACCTCAGCCCAGGAGGTCACCCAGCGCCCCCAGAGGCTCCTGGGA 1132
Db    357 GGAGCTGGGGGGATCCCCACCTCAGCCCAGGAGGTCACCCAGCGCCCCCAGAGGCTCCTGGGA 416

Qy   1133 CCCAATGGCCTCTGGATGAAACGTCAGGAATTACTACACAGAAAATCAGAGAAAGAGAGA 1192
Db    417 CCCAATGGCCTCTGGATGAAACGTCAGGAATTACTACACAGAAAATCAGAGAAAGAGAGA 476
```

```
RESULT 4
BQ072636

LOCUS       BQ072636                1055 bp    mRNA    linear    EST 02-APR-2002
DEFINITION  AGENCOURT_6763126 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5756099
            5', mRNA sequence.
ACCESSION   BQ072636
VERSION     BQ072636.1  GI:19901682
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1055)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1296  row: c  column: 12
            High quality sequence stop: 692.

FEATURES             Location/Qualifiers
     source          1..1055
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5756099"
                     /tissue_type="leukocyte"
                     /clone_lib="NIH_MGC_118"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     /clone_lib="NIH_MGC_118 RNA source leukocytes from anonymous pool of
                     non-activated adult donors. Library is oligo-dT primed
                     and directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.7 Kb, insert size range
                     1.2-3.3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."

BASE COUNT      220 a      349 c      282 g      202 t      2 others
ORIGIN

Query Match  15.1%;   Score 819.8;   DB 13;   Length 1055;
Best Local Similarity  98.7%;   Pred. No. 2.1e-100;
Matches  846;  Conservative  0;  Mismatches  9;  Indels  2;  Gaps  2;

Qy      1 GCCCCAAGGGCCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGCGCCAGCCCCGGGGCTCCACTCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     50 GCCCCAAGGGCCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGCAGCGCCCGGGGCTCCACTCT 109

Qy     61 GGGGTTCTGAAAGCCCATTCCCTGCTCTCGCGGCTCCTCCCACACCCCACTCTTCTTCAGCCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    110 GGGGTTCTGAAAGCCCATTCCCTGCTCTCGCGGCTCCTCCCACCCCACTCTTCTTCAGCCTT 169

Qy    121 GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGAGTCCAGGAGAGAGGAAGAAATCTGAGGAACA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    170 GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGAGAGGAAGAAATCTGAGGAACA 229

Qy    181 CAGAACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCACATCTCCCCTCACCCCTCAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    230 CAGAACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCACATCTCCCCTCACCCCTCAC 289

Qy    241 CCTCCCTGCCTGGCCCTGGACCCCCATCCCAGGAGACCTCCTATCAGCTGACTTCTTCCAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    290 CCTCCCTGCCTGGCCCTGGACCCCCATCCCAGGAGACCTCCTATCAGCTGACTTCTTCCAGT 349

Qy    301 GTCTTGCAGGCCCCTCTGGGCTCTCCCCTGGCTTTTCCTACCACTCCCCCTCTAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    350 GTCTTGCAGGCCCCTCTGGGCTCTCCCCTGGCTTTTCCTACCACTCCCCCTCTAT 409

Qy    361 CGGCGGTCTATCTGTAGGTGCCCCTGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    410 CGGCGGTCTATCTGTAGGTGCCCCTGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA 469

Qy    421 GACGGTAAGAGCCAAGGCAAAAGGACAGCACTGTTCTCTGCCTGATACCCTCCACCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    470 GACGGTAAGAGCCAAGGCAAAAGGACAGCACTGTTCTCTGCCTGATACCCTCCACCAC 529

Qy    481 CTGGGAACATCCCCCAGACACCCTCTTAACTCCCGGGACAGAGATGGCTGGCGGAGCCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    530 CTGGGAACATCCCCCAGACACCCTCTTAACTCCCGGGACAGAGATGGCTGGCGGAGCCTGG 589

Qy    541 GGGCGCGCCTCGGCCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAGGAGTTCCAGCTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    590 GGGCGCGCCTCGGCCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAGGAGTTCCAGCTT 649

Qy    601 CTGCTCCGCCAATAAAGGCGCACTCCAGGAGAGCTCTTCGGGTGAGACACCCGCTCAGCCGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    650 CTGCTCCGCCAATAAAGGCGCACTCCAGGAGAGCTCTTCGGGTGAGACACCCGCTCAGCGAG 709

Qy    661 AAGACGAGTGGCATGGAGAGTGGCCTTCGGTACCTGTGGCCTCAGTATGGTGGGACCACGGGGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    710 AAGACGAGTGGCATGGAGAGTGGCCTTCGGTACCTGTGGCCTCAGTATGGTGGGACCACGGGGCC 769

Qy    721 TGGGACCTTAGCCCTTCCATACCTGGGAGCCAGCAGGATGGGGCTCACTGTGTCGCGCCC-AAGC 779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db    770 TGGGACCTTAGCCCTTCCATACCTGGGAGCCAGCAGGATGGGGCTCACTGTGTCGCGCCCAAAGC 829
```

```
Qy    780  CCAGGAAGGGGCAGGGCCACTCTCTCCCTCATTCCCTCATTCCCTACAGCCCAAGTGAACCCCACCT-GG  838
           ||||||||||||||||| ||||||||||||||||||||||||  ||||||||||||||||||||||| ||
Db    830  CAGGGAAAGGGCAGGGCCACTCTTCCTCATTCCCTAAAGCCCAAGTGACCCCACCTGGG  889

Qy    839  GGTCTCCCAGCCAACCC  855
           |||||||||||||||||
Db    890  GGTCTCCCAGCCAACCC  906
```

RESULT 5

LOCUS   BM549709   1046 bp   mRNA   linear

DEFINITION AGENCOURT_6544160 NIH_MGC_118 Homo sapiens cDNA clone   EST 20-FEB-2002
            IMAGE:5745766 5', mRNA sequence.

ACCESSION   BM549709

VERSION   BM549709.1 GI:18785339

KEYWORDS   EST.

SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 1046)

AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE   Unpublished

JOURNAL

COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12769 row: d column: 23
            High quality sequence start: 14
            High quality sequence stop: 649.

FEATURES   Location/Qualifiers
     source   1..1046
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5745766"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_118"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated adult donors. Library is oligo-dT primed
                     and directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 027. Note:
                     this is a NIH MGC Library."

BASE COUNT   227 a   293 c   295 g   231 t

ORIGIN

Query Match   15.0%; Score 814.4; DB 12; Length 1046;
Best Local Similarity 98.4%; Pred. No. 1.1e-99;
Matches   844; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

```
Qy   3712  TCTCAAGGGACACTTGCATACGAAGCCCTTTGGGGACTGACGGATGACTTCTGGGGGCCCACG  3771
           ||||||||| |||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     20  TCTCAAGGGGAACCTTGCATACGAAATCCTTTGGGGACTGACGGATGACTTCTGGGGGCCCACG  79

Qy   3772  GGGCCTGTGTGCTACTTGAGGTAGTTGCACAAAGAAAAAGAAACTTGTACCGAGTTCACTTCCCT  3831
           |||||||||| ||||| |||||||||| |||||||||||||   |||||||||||||||||||||
Db     80  GGGGCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAAGAACTTGTACCGAGTTCACTTCCCT  139

Qy   3832  GTAGCTGGCTGGCTCCTACACGTCCCAACACGGGGTCTCGCTTGTTGTGATGAGGAAGCGGTG  3891
           ||||||||| ||||||||||| ||||||||| ||||||| |||||||||||||||||||||||
Db    140  GTAGCTGGGCTCCTACACGCTGGGCCCAACACAGGGGTCTCTGCTTGTTGTGATGAGGAAGCGGTG  199
```

```
Qy   3892  ACGGTTGAGATTGAATTCGTGTGTGTGTGGGGGGACCCAGTTCCTGGGGTGAGATCAACCCACAGCAC  3951
           |||||||||| |||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db    200  ACGGTTGAGAATTGAATTCGTGTGTGTGGGGACCCAGTTCCTGGGGTGAGATCAACCCACAGCAC  259

Qy   3952  AGCTGGATGGTGGCAGGGGCCTTCGTGGGACATCAAGGCTGAGCCTGGAGCTGTGTGGAAGCT  4011
           |||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db    260  AGCTGGATGGTGGCAGGGGCCTTCGTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGTGGAAGCT  319

Qy   4012  GTGCACCTCCCTCACTTTGTGGCTCTCCAAGGGGGCCCATGTGGCACATCCCTGTTCCAA  4071
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    320  GTGCACCTCCCTCACTTTGTGGCTCTCCAAGGGGGCCCATGTGGACACATCCCTGTTCCAA  379

Qy   4072  ATGGGCCCACTTTAAAGAGAGGAGGGGGGGATGCTCCTGGAGAAGCCAGGGGTGGAGCTGCAT  4131
           |||||||||||||||||| |||||||||| ||||||||||||||||||||   |||||||||||
Db    380  ATGGGCCCACTTTAAAGAGGAGGGGGGATGCTCCTGGAGAAGCCAGCCAGGGGTGGAGCTGCAT  439

Qy   4132  CACATAGTTCTGGAAAAACCCCAGCGTTCTCCCCCTTGGGAGTCCTCCTGAAAATGATCCAT  4191
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    440  CACATAGTTCTGGAAAAACCCCAGCGTTCTCCCCCTTGGGAGTCCTCCTGAAAATGATCCAT  499

Qy   4192  AATGCCCTGGGCCTTCATTCCCGTCACCTCGTGTGGTTGTTGCTTTACCACCGGGTCATCCT  4251
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500  AATGCCCTGGGCCTTCATTCCCGTCACCTCGTGTGGTTGTTGCTTTACCACCGGGTCATCCT  559

Qy   4252  GAGGGAAGTCACCTTCCACCTCACCTCACTGATCCAAGTGACTGCTCCATTCGGAGGAGGAACTG  4311
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560  GAGGGAAGTCACCTTCCACCTCACCTCACTGATCCAAGTGACTGCTCCATTCGGAGGAGGAACTG  619

Qy   4312  GAGCCTCTGCTATCGAAGCCCTGGAGAAGACCAGCGTTCTCGGAGGTTCTACGGTTGGCCAC  4371
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    620  GAGCCTCTGCTATCGGAGCCCTGGAGAAGACCAGCGTTCTCGGAGGTTCTACGGTTGGCCAC  679

Qy   4372  TTGGGATCAGGGATCAGGGCTGCAAGTGAAAGACAAGAAAGAGATGAGACTGCTGGT-GTGGGA  4430
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||||
Db    680  TTGGGATCAGGGATCAGGGCTGCAAGTGAAAGACAAGAAAGAGATGAGACTCTGGTGGTGGGA  739

Qy   4431  GGGCCTTGGGTGAAAACCAGGAGAGTCTCATGCCTGCAACTACTCTGATCCCTCCCAGCCCGCAT  4490
           |||||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||
Db    740  GGGCCTTGGGTGAAAAACAGGAGAGTCTCATGCCTGCAACTACTCTGATCCCTCCAGCCCGCAT  799

Qy   4491  AGCCGTACCTTCACCTCTGATGCCCCGCAGTTGC--TGCACTTGTTGGACCAGTATCGA  4548
           ||||||||||||||||||||||||||||||||||  |||||||| |||||||||||||
Db    800  AGCCGTACCTTCACCTCTGATGCCCCGCAGTTGCCCAGTTGCCCACTTGGGGGACCAGTATCGA  859

Qy   4549  GAGCAGCTGATTAGCCCGA  4566
           |||||||||||||||||||
Db    860  GAGCAGCTGATTAGCCCCA  877
```

RESULT 6

LOCUS   BX457482   1110 bp   mRNA   linear

DEFINITION BX457482 Homo sapiens THYMUS Homo sapiens cDNA clone   EST 22-MAY-2003
            CS0CAP008YN15 5-PRIME, mRNA sequence.

ACCESSION   BX457482

VERSION   BX457482.1 GI:31028889

KEYWORDS   EST.

SOURCE   Homo sapiens (human)

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 1110)

AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE   Full-length cDNA libraries and normalization

JOURNAL   Unpublished

COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8396.r For
            more information about this cluster, see

```
                http://www.genoscope.cns.fr/
                cgi-bin/cluster.cgi?seq=CS0CAP008CG08QP1&cluster=8396.r. Contact :
                Feng Liang Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
                Faraday Avenue Genoscope sequence ID : CS0CAP008CG08QP1.

FEATURES        Location/Qualifiers
    source      1..1110
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0CAP008YN15"
                /tissue_type="THYMUS"
                /clone_lib="Homo sapiens THYMUS"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a Not1-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."

BASE COUNT      218 a      270 c      257 g      262 t      103 others
ORIGIN

Query Match         14.5%;   Score 787.4; DB 13; Length 1110;
Best Local Similarity   88.3%;   Pred. No. 4.6e-96;
Matches 848; Conservative   58; Mismatches   48;  Indels   6;  Gaps   6;

Qy   1644 GTCCAAGGTGGTGAGTCTTCCTGCTGAGGTCTCGAGGCTCATCGGGAAAAAGAGATGGGACAGCAGCCACTCCGGCTCC 1703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 GACCCACGCGTCCGGGTCTCGCTCAGCCTCATCGGGAAAAAGAGATGGGACAGCAGCCACTCCGGCTCC 131

Qy   1704 CATTAGAGACAGATCCTGTCTAGGCGGCCGCAGAGGCGGTGCTTGCTTCTTCTTCATCCTCGATGGGTGTAGAGTGA 1763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 CATTAGAGACAGATCCTGTCTAGGCGGCCGCAGAGGCGGTGCTTGCTTCTTCTTCATCCTCGATGGGTGTAGAGTGA 191

Qy   1764 GCCAGGAGGATGGGGTCTTTGCAGGAGCCGAGTTCTGAGCTCTGTCTCTGCCACTGGAGCTCAGCCACA 1823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 GCCAGGGATGGGGTCTTGCAGGAGGCCGAGTTCTGAGCTCTGTCTCTGCCACTGGAGCCCAGCCACA 251

Qy   1824 GCGGCGGCGGATGCACTGCTGGGCCAGTTGCTGGGCGGAAAACTATATCTTCCCGAGGCGGCCATCCTT 1883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 GCGGCGGCGGATGCACTGCTGGGCCAGTTGCTGGGCGGAAAACTATATCTTCCCGAGGCGGCCATCCTT 311

Qy   1884 CCTGATCACGGCGTCGGACCACAGCTTCTGCAGAACCTCAATTCCTTCTTTGGACAGGCACG 1943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 CCTGATCACGGCGTCGGACCACAGCTTCTGCAGAACCTCAATTCCTTCTTTGGACAGGCACG 371

Qy   1944 TTTGGGTAGGAGGTCCTGGGGTTCTCTGAGTCCAGCAGGGAGGAAGAATAATTTCTACAGATATTT 2003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 TTTGGGTAGGAGGTCCTGGGGTTCTCTGAGTCCAGCAGGGAGGAATATATTTCTA CAGATATTT 431

Qy   2004 CACAGATGAAAGGCAAGCAATTAGAGCCTTTAGGTTGGTGGTCAAATCAACAAAAGAGCTCTG 2063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 CACAGATGAAAGGCAAGCAATTAGAGCCTTTAGGTTGGTGGTCAAATCAACAAAAGAGCTCTG 491

Qy   2064 GGCCCTGTGTCTTGTGCCCTGGGTGTCCTGGCTGCCTGGCGCAGTTGCCTGCCGATGCAGCAGAT 2123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 GGCCCTGTGTCTTGTGCCCTGGGTGTCCTGGCTGCCTGGCGCAGTTGCCTGCCGATGCAGCAGAT 551

Qy   2124 GAAGCGGAAGGAAAAAACTCACACTGACTTCCAAGACCACACAACCCTCTGTCTACATTA 2183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 GAAGCGGAAGGAAAAAACTCACACTGACTTCCAAGACCACACAACCCTCTGTCTACATTA 611

Qy   2184 CCTTGCCCCAGGCTTCTCCAAGCTCAGCCATTGGGGACCCAGCTCAGAGA--CCTCTTGCCCTCTC 2242
          ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db    552 GAACCGGAAGGAAAAAACTGACTTCCAAGACCACACACCCTCTGTCTACATTA 611

Qy   2243 TGGGCTGCTGAGGGCATCTGGCAAAAA--AAGACCCTTTCAGTCCAGATGACCTCAGGAG 2301
          |||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db    672 TGGGCTGCTGAGGGCATCTGGCAAAAASAAGACCCTTTTCAGTCCAGATGACCTCAGGAAG 731

Qy   2302 CATGGG--TTAGATGGGGGCCATCATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCA 2360
          |||||   ||||||||||:|||||||||||||||||||||||||||||||:||||||||||
Db    732 CATGGGGTTTAGATGGGGGGGSCATCATCATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCA 791
```

```
Qy   2361 CCCCATCCCTCTGAGCCTACAGCTTCATTCACCTCTGTTTCACCTCGTTTCCCAAGAGTT-CTTTGCAGCGAA 2419
          |||||||||||||||||||||||||:||||:::||||||||||||||||||::|||||||||||   ::|||||:
Db    792 CCCCATCCCTCTGAGCCTACAGCTTYTTMTTMCCTCTGTTTCCMMGAGTTYYTTGSMGCAM 851

Qy   2420 TGTCCTATGTCTTGGAGGATGAGAAGAGGGGAAGGGAGGTAAACATTCATCATAGA-T 2478
          |||||||||||||||||||||:||||||:|::||||:||||||||||||:::|||||:
Db    852 TGTCCTMGTCTTGGAGGATGAGAMGGGGACHMCMYKGGTNKWTTKGGGAMYAMCCWMMMCG 911

Qy   2479 TTCGAAAAGACACGCTAGA-AGCATATGGAATACATKGGCCTGTTTGGGGGCATCAACCACACAG 2537
          ||||||||||||||||   :||||:|::|::|||||||:::|:|::||   :|:|||||||||||
Db    912 TTGGMACCCGMCKSTCSMCCGMCACTGGMCHMCMYKGGGTNKWTTKGGGAMYAMCCWMMMCG 971

Qy   2538 TTTCCTTATTGGGCCTGTTTAAGTGATGAGGGGGAGAGAGAGGAAGACATCTTTTCACTG 2597
          |||   ||||||||:|||||||||||||||||:||||:|||||||:||||||:||||||
Db    972 TTTGATMCWKGGGCTTTWMCVTWCWKASKGKGWGCACMCMTMKYCTMCACYTMKAMCKKTG 1031
```

```
FEATURES        Location/Qualifiers
    source      1..768
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5211072"
                /tissue_type="leukocyte"
                /clone_lib="NIH_MGC_118"
                /lab_host="DH10B"
                /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                (destroyed); RNA source leukocytes from anonymous pool of
                non-activated adult donors. Library is oligo-dT primed
                and directionally cloned (EcoRV site is destroyed upon
                cloning). Average insert size 1.7 kb, insert size range
                1.2-3.3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 027. Note:
                this is a NIH_MGC Library."

BASE COUNT      176 a      213 c      225 g      154 t
ORIGIN

Query Match         14.1%;   Score 767; DB 12; Length 768;
Best Local Similarity   100.0%;   Pred. No. 2.7e-93;
Matches 767; Conservative   0; Mismatches   0;  Indels   0;  Gaps   0;

Qy   3119 CCCTGACCGAGCTGGACTTCAATGTGCTCACGGATGCTGGAGCCAAAACACCTTT 3178
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCCTGACCGAGCTGGACTTCAATGTGCTCACGGATGCTGGAGCCAAAACACCTTT 60
```

```
Qy   3179  GCCAGAGACTGAGACAGCCGAGCTCAAGCTACACGGACTGGTCAGCTGTGGCC  3238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GCCAGAGACTGAGACAGCCGAGCTCAAGCTACACGGACTGCAGCTGTGTGGCC  120

Qy   3239  TCACGTCTGCTGCTGCCAGGA... ...CAGGCCTGAAGG  3298
           |||||||||||||||||||||
Db    121  TCACGTCTGCTGCTGCCAGGA... ...CAGGCCTGAAGG  180

Qy   3299  AGCTAGACCTGCAGCAGAACAACTGGA... ...GTTGGGGGC  3358
Db    181  AGCTAGACCTGCAGCAGAACAACTGGA... ...GTTGGGGGC  240

Qy   3359  TCAGGCATCCTGCCTGCAAACT...  3418
Db    241  TCAGGCATCCTGCCTGCAAACT...  300

Qy   3419  AGATGAGGCAGGAACTGAGGGGCCCT...  3478
Db    301  AGATGAGGCAGGAACTGAGGGGCCCT...  360

Qy   3479  GACGGAAACCAAGTGTGATGAACCCTA...  3538
Db    361  GACGGAAACCAAGTGTGATGAACCCTA...  420

Qy   3539  GCACATCCTCACTCAAGCGCAGAGA...  3598
Db    421  GCACATCCTCACTCAAGCGGCAGA...  480

Qy   3599  AGGCTAAATCTCAAACTCTCTGGAC...  3658
Db    481  AGGCTAAATCTCAAACTCTCTGGAC...  540

Qy   3659  AAAGCTCCCAGAGAGGTAGTACCGG...  3718
Db    541  AAAGCTCCCAGAGAGGTAGTACCGG...  600

Qy   3719  GGGACCTGCATACAGGAAGCCTTT...  3778
Db    601  GGGACCTGCATACAGGAAGCCTTT...  660

Qy   3779  TGGGCTACTGAGGGTAGTTGACAA...  3838
Db    661  TGGGCTACTGAGGGTAGTTGACAA...  720

Qy   3839  GCTCCTACCGCTGGCCCAACACGG...  3885
           ||||||||||||||||||||||||
Db    721  GCTCCTACCGCTGGCCCAACACGG...  767
```

LOCUS   CD515485   894 bp   mRNA   linear   EST 06-JUN-2003
DEFINITION AGENCOURT 14364625 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30407209 5', mRNA sequence.
ACCESSION   CD515485
VERSION   CD515485.1 GI:31447203
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
  COMMENT   Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

RESULT 8
CD515485
CD515485

FEATURES   Location/Qualifiers
     source   1..894
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30407209"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
                     (destroyed); library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH MGC Library."

BASE COUNT   194 a   241 c   268 g   189 t   2 others
ORIGIN

Query Match   13.8%;   Score 749.4; DB 14; Length 894;
Best Local Similarity 95.7%; Pred. No. 6e-91;
Matches 780; Conservative   0; Mismatches 33; Indels   2; Gaps   1;

```
Qy   3202  TGCAAGCTACACGCGACTGCAAGCTGGTCAGCTGTGTGGCCTCACGTCTGCTGCTGCCAGGAC  3261
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TGCAAGCTACACGCGACTGCAAGCTGGTCAGCTGTGTGGCCTCACGTCTGCTGCTGCCAGGAC  60

Qy   3262  CTGGGCCTCTGTGCTTAGTGCCAGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAACAAC  3321
Db     61  CTGGGCCTCTGTGCTTAGTGCCAGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAACAAC  120

Qy   3322  CTGGGATGACCGTTGGGCGTGCCAGCTGCTCTGTCAGGGGTCAGGCATCCTGCCTGCAAACTC  3381
Db    121  CTGGGATGACCGTTGGGCGTGCCAGCTGCTCTGTCAGGGGTCAGGCATCCTGCCTGCAAACTC  180

Qy   3382  ATACCGCCTGGGGCTGGACCAGACAACTCTGAGTGATGAGATGAGGGCC  3441
Db    181  ATACCGCCTGGGGCTGGACCAGACAACTCTGAGTGATGAGATGAGGGCC  240

Qy   3442  CTGGAGCCAGGAGAAACCTCAGCTGTCATCTTCAGCAGTGATGACC  3501
Db    241  CTGGAGCCAGGAGAAACCTCAGCTGTCATCTTCAGCAGTGATGACC  300

Qy   3502  CCTACTGAGGGGCCTGGATACGGGAGAGATGAGTAATAGCACATCCTCACTCAAGCGCAG  3561
Db    301  CCTACTGAGGGGCCTGGATACGGGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAG  360

Qy   3562  AGACTCGGATCAGAGAGGCGGGCCTTCCCATGTTGCTCAAGGCTAAATCTCAAACTCTGGAC  3621
Db    361  AGACTCGGATCAGAGAGGCGGGCCTTCCCATGTTGCTCAAGGCTAAATCTCAAACTCTGGAC  420

Qy   3622  GTGAGGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAAAGCTCCCAGAGAGGTAGTACCG  3681
Db    421  GTGAGGCAAGATCTTCCCAATTGCTGAGATTGCAGAGGAAAGCTCCCAGAGAGGTAGTACCG  480

Qy   3682  GTGGAACTCGTTGTGCGTGCCTTCGCTGCCTTCTCCTGCCTTCTCCTGCCTTTG  3741
Db    481  GTGGAACTCTTGTGCGTGCCTTCGCTGCCTTCTCCTGCCTTCTCCTGCCTTTG  540

Qy   3742  GGGGACTGACGATGACTTCTGGGGCCCCACGGGCCCTGTGGCCTACTGAGGTAGTTGACAAA  3801
Db    541  GGGGACTGACGATGACTTCTGGGGCCCCACGGGCCCTGTGGCCTACTGAGGTAGTTGACAAA  600

Qy   3802  GAAAAGAAACTTGTACCGAGTTCACTTCCCTGTAGCTGGGCTCCTACCGCTGGCCCAACACG  3861
Db    601  GAAAAGAAACTTGTACCGAGTTCACTTCCCTGTAGCTGGGCTCCTACCGCTGGCCCAACACG  660

Qy   3862  GGTCTCTGCTTTGTGATGAGAGAAGGGTGACACGGTTGAGATTGAATTCTGTGTGTGGGAC  3921
Db    661  GGTCTCTGCTTTGTGATGAGAGAAGGGTGACACGGTTGAGATTGAATTCTGTGTGTGGGAC  720
```

Left column:

```
Db    661  GGTTCTCTGCTTTGTTGATGATGAGAGAAGAGGTTGACCGGTTGAGATTGAATTCTGTGTGTGTGGAC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3922  CAGTT---CCTGGGTGAGATCAACCACCAGCACAGCTGATGAGTGCAGGCCTTCTGCTGTGG  3979
           || ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CAGTTTCCTGGGGTGAGATCAACCACAGCACANCTGGGATGGTGGCAGGGCCCTCTTGC  780
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy   3980  ACATCAAGGGCTGAGCCTGGAGCTGTGTGAAGCTGTG  4014
           ||||||||||||||||||||||||||||||||||||
Db    781  TGGNACATCAAGGGCTTGAACCCTGGGGAAGCTGTG  815
```

RESULT 9

LOCUS　　　　BG682080

DEFINITION　602630057F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754607 5',
　　　　　　　mRNA sequence.

ACCESSION　　BG682080

VERSION　　　BG682080.1　GI:13913477

KEYWORDS　　EST.

SOURCE　　　Homo sapiens (human)

　ORGANISM　Homo sapiens
　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE　　1　(bases 1 to 791)

AUTHORS　　NIH-MGC http://mgc.nci.nih.gov/).

TITLE　　　National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL　　Unpublished

COMMENT　　Contact: Robert Strausberg, Ph.D.
　　　　　　Email: cgapbs-r@mail.nih.gov
　　　　　　Tissue Procurement: James Cleaver, M.D.
　　　　　　cDNA Library Preparation: Life Technologies, Inc.
　　　　　　cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
　　　　　　DNA Sequencing by: Incyte Genomics, Inc.
　　　　　　Clone distribution: MGC clone distribution information can be
　　　　　　found through the I.M.A.G.E. Consortium/LLNL at:
　　　　　　http://image.llnl.gov
　　　　　　Plate: LLAM10616　row: b　column: 16
　　　　　　High quality sequence stop: 770.

FEATURES　　　　　　　　　Location/Qualifiers
　　source　　　　　　　　1..791
　　　　　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　　　　　/clone="IMAGE:4754607"
　　　　　　　　　　　　　/tissue_type="squamous cell carcinoma"
　　　　　　　　　　　　　/lab_host="DH10B (T1 phage-resistant)"
　　　　　　　　　　　　　/clone_lib="NCI_CGAP_Skn4"
　　　　　　　　　　　　　/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
　　　　　　　　　　　　　Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
　　　　　　　　　　　　　Average insert size 1.5kb. Library constructed by Life
　　　　　　　　　　　　　Technologies.Note: this is a NCI_CGAP Library."

BASE COUNT　　　195 a　　203 c　　200 g　　193 t

ORIGIN

```
Qy   2045  AATCAAACAAAAGAGCTCTGGGCCCTGGTCGTGTCTGTGGCCTGGCCTGGCA  2104
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3  AATCACACAAAAGAGCTCTGGGCGCCCCTGTGTCTTGTGCCCTGGGTGTCTGGGCTGGG-CTGCA  61
```

Query Match　　13.4%;　　Score 727;　　DB 10;　　Length 791;
Best Local Similarity　98.2%;　Pred. No. 6.1e-88;
Matches 777;　　Conservative　0;　Mismatches 10;　Indels 4;　Gaps 4;

```
Qy   2105  CTTGCCTGATGCAGCGATGAACGGAGGGAAAAAAACTCACACTGACTTCCAAGACACCACCA  2164
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62  CTTGCCTGATGCAGCAGATGAACGGAGGGAAGGAAAAAAACTCACACTGACTTCCAAGACACCA  121
Qy   2165  CAACCCCTTCTGTCTTACATTACCTTGCCCCAGGCCTCAGCCAGTGGGACCCCCAGC  2224
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122  CAACCCCTTCTGTTCTACATTACCTTGCCCAGGCTCAGCCATGGGACCCCAGC  181
Qy   2225  TCAGAGACCTCTGCTCTCTGGCGTGCTGAGGGGCCATCTGGCAAAAAAGACCCCTTTTCAGTC  2284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column:

```
Db    182  TCAGAGACCTCTCGTCTCTCTGGGCTGCTGAGGGCATCTGGCCAAAAAAAGACCCCTTTCAGTC  241
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2285  CAGATGACCTCAGGAAAGCAAGGGGTTAGATTGAGTGGGGCCATCATCTCCACCTTCTTGAAGATGG  2344
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242  CAGATGACCTCAGGAAAGCAGTTAGATGGGGCCCATCATCTCCACCTTCTTGAAGATGG  301
Qy   2345  GTATTCTTCAAGACGCACCCCATCCCTCTGAGCTACAGCTTCATTCACCTCGTTTCCAAG  2404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302  GTATTCTTCAAGACGCACCCCATCCCTCTGAGCTACAGCTTCATTCACCCTCGTTTCCAAG  361
Qy   2405  AGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGATGAGAAGGGGGGAGGAGGTAAACATTCTA  2464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362  AGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGATGAGAAGGGGGAGGAGGTAAACATTCTA  421
Qy   2465  ATTGCATCAATAGATTTGGAAAAGACGCTAGAACGCATATGAATACATCATGCCTGTTTGGGG  2524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422  ATTGCATCAATAGATTTGGAAAAGACGCTAGAAGGGCCCATATGAATACATGGCCTGTTTGGGG  481
Qy   2525  CATCAACCACACACGTTTTCCTATTGGGCCCTGTTTAAGTTGATGTCAGTGGAGGGAGA  2584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482  CATCAACCACACGTTTTCCTATTGGGCCCTGTTTAAGTTGATGTCAGTGGAGGGGAGA  541
Qy   2585  ACATCTTTCACTGCCGGCCTGTCTTCCAGGGGAGAGAACCTGAGTGCAGTGAGTGGTTCCCGTTCCCTGC  2644
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542  ACATCTTTCACTGCCGGCTGTCTTCAGGGGAGAGAACCTGAGTGCAGTGAGTGGGTTCCCGTTCCCTGC  601
Qy   2645  AGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCCTCCACTGCTTGTGTACGGAGACTCGGAACA  2704
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602  AGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCCTCCACTGCTTGTGTACGGAGACTCGGAACA  661
Qy   2705  AAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAAATGGGCA-TGTGTGTAGAAACA  2763
           |||||||||||||||||||||||||||||||||||||||||||||||| |
Db    662  AAA-GTTCCTGACACAGTGATGGCCCCATTTTCGAAGAAAATGGGCCATTTGTGTGTGAGAAACA  720
Qy   2764  GACATGGAGCTCTTAGTGTGCACTTTCTTGGCATT-AAATTCAGCCGCCGCCAGGTGAAGAAGCT  2822
           ||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    721  GACATGGAGCTCTTAGTGTGCACTTTCTTGGCATTAAAATTCAGCCGCCCGCCAGTGAAGAAGCT  780
Qy   2823  TCAGCTGATTG  2833
           |||||||||||
Db    781  TCAGCTGATTG  791
```

RESULT 10

LOCUS　　　　BQ576345/c

DEFINITION　UI-H-EZl-bbi-n-10-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
　　　　　　　UI-H-EZl-bbi-n-10-0-UI 3', mRNA sequence.

ACCESSION　　BQ576345

VERSION　　　BQ576345.1　GI:21479662

KEYWORDS　　EST.

SOURCE　　　Homo sapiens (human)

　ORGANISM　Homo sapiens
　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE　　1　(bases 1 to 764)

AUTHORS　　NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE　　　National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
　　　　　　Tumor Gene Index

JOURNAL　　Unpublished

COMMENT　　Contact: Robert Strausberg, Ph.D.
　　　　　　Email: cgapbs-r@mail.nih.gov
　　　　　　Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
　　　　　　Orthopaedics
　　　　　　cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
　　　　　　cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
　　　　　　DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
　　　　　　Clone Distribution: Clone distribution information can be obtained
　　　　　　from Dr. M. Bento Soares, bento-soares@uiowa.edu
　　　　　　The following repetitive elements were found in this cDNA
　　　　　　sequence: 1-71, >POLY_A#simple repeat (matched compliment) 72-100,
　　　　　　>AT_rich#low_complexity 157-242, >MER41B#LTR/MER4-group
　　　　　　Seq primer: M13 FORWARD

FEATURES   Location/Qualifiers
source   1..764
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-H-EZ1-bbi-n-10-0-UI"
                  /tissue_type="Chondrosarcoma Grade II"
                  /dev_stage="Adult"
                  /lab_host="DH10B (Life Technologies)"
                  /clone_lib="NCI_CGAP_Ch2"
POLYA=Yes.
                  /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
                  with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
                  NCI_CGAP_Ch2 is a normalized cDNA library containing the
                  following tissue(s): Chondrosarcoma Grade II. The library
                  was constructed according to Bonaldo, Lennon and Soares,
                  Genome Research, 6:791-806, 1996. First strand cDNA
                  synthesis was primed with an oligo-dT primer containing a
                  Not I site. Double stranded cDNA was ligated to an EcoR I
                  adaptor, digested with Not I, and cloned directionally
                  into pT7T3-Pac vector. The oligonucleotide used to prime
                  the synthesis of first-strand cDNA contains a library tag
                  sequence that is located between the Not I site and the
                  (dT)18 tail. The sequence tag for this library is
                  TGATCACGGT.
                  TAG_LIB=UI-H-EZ1
                  TAG_TISSUE=grade-2-chondrosarcoma
                  TAG_SEQ=ATCTAATATG"

BASE COUNT   173 a   175 c   239 t   2 others
ORIGIN   175 g

Query Match   13.3%;   Score 722.4; DB 13;   Length 764;
Best Local Similarity   99.5%;   Pred. No. 2.5e-87;
Matches 734; Conservative   0; Mismatches   3; Indels   1; Gaps   1;

Qy   4702 CGGAAGTGCAAAAGATGGAACTCTACCAAAGCCCCTGAAGGAGGAGACCCATCCTCACCTCATTATG 4761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   754 CGGAAGTGCAAAAGAT-GACTCTACCAAAGCCCCTGAAGGAGGAGACCCATCCTCACCTCATTATG 696

Qy   4762 GAACTTCTGGGAGAAGGGCAGCAAAAAAGGAGACTCCTGCCACTCAGCCAGCTGAAAGTAATCAAC 4821
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 GAACTTCTGNGAGAAGGGCAGCAAAAAAGNGACTCCTGCCACTCAGCCAGCTGAAAGTAATCAAC 636

Qy   4822 ACCAGCCCTTGACCCTTGAGTCCTGGCCTTTGGGCCTTTGGGTCTCAGTTTTCTT 4881
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   635 ACCAGCCCTTGACCCTTGAGTCCTGGCCTTTGGGCCTTTGGGTCTCAGTTTTCTT 576

Qy   4882 TCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAACTTTGAT 4941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   575 TCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAACTTTGAT 516

Qy   4942 GATGCCTTTGCTNGGGCATTAATGTNGTTCCATGCCACACGGGGGCCCCAGTCCAG 5001
           ||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db   515 GATGCCTTTGCTGGGGCATTAATGTGTCCATGCCACACGGGGGCCCCAGTCCAG 456

Qy   5002 GTGGGCCTAACAGCATCTCAGGGAATGTTCCATCTGGAGCTGGCAAGACCCTGCCAGACCTC 5061
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455 GTGGGCCTAACAGCATCTCAGGGAAATGTCCATCTGGAGCTGGCAAGACCCTGCCAGACCTC 396

Qy   5062 ATAGAGCCTCATCTGGTTGGGCCACAGCAGACCTTAGAGCCCTCCGGGATCCCATCCAGG 5121
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   395 ATAGAGCCTCATCTGGTTGGGCCACAGCAGACCTTAGAGCCCTCCGGATCCCATCCAGG 336

Qy   5122 CGCAAAGAGAGAATAGGAGGGACATGGAACCATTTGCCTCTGTGTCACAGGGTGAGC 5181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   335 CGCAAAGAGAGAATAGGAGGGACATGGAACCATTTGCCTCTGTGTCACAGGGTGAGC 276

Qy   5182 CCCAAAAATTGGGGTTCAGGGTGGGGGAGGGGAGGCCCACGTGATGGGGTGAGGAAGATC 5241
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275 CCCAAAAATTGGGGTTCAGGGTGGGGAGGCCACGTGGAATTCCTTTGTACAGGAAGATC 216

Qy   5242 TACAAGAGCAAGCCAACAGAGTAAAGTGGAAGGAAGTTTATTCAGAAAAATAAAGGAGTAT 5301
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   215 TACAAGAGCAAGCCAACAGAGTAAAGTGGAAGGAAGTTTATTCAGAAAAATAAAGGAGTAT 156

Qy   5302 CACAGCGTCTTTTAGAAATTTGTCTAGCAGGGCTTTCCAGTTTTTTACCAGAAAACCCCTATAA 5361
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   155 CACAGCGTCTTTTAGAAATTTGTCTAGCAGGGCTTTCCAGTTTTTTACCAGAAAACCCCTATAA 96

Qy   5362 ATTAAAAATTTTTTACTTAAATTTAAGAAGTTAAAATTACAAAAAAGAAAAAAATGAAAA 5421
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   95 ATTAAAAATTTTTTACTTAAATTTAAGAAATTAAAATTACAAAAAAGAAAAAAATGAAAA 36

Qy   5422 TAAAGGAATAAGAAAGTTA 5439
           |||||||||||||||||||
Db   35 TAAAGGAATAAGAAAGTTA 18

RESULT 11
CA308213/c

LOCUS   CA308213   686 bp   mRNA   linear   EST 01-NOV-2002
DEFINITION   UI-H-FT1-bib-1-22-0-UI.s1 NCI_CGAP_FT1 Homo sapiens cDNA clone
                 UI-H-FT1-bib-1-22-0-UI 3', mRNA sequence.
ACCESSION   CA308213
VERSION   CA308213.1   GI:24471267
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 686)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                 Tumor Gene Index
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
                 Email: cgapbs-r@mail.nih.gov
                 Tissue Procurement: Dr. Gary W. Hunninghake, U of I
                 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                 Clone Distribution: Clone distribution information can be obtained
                 from Dr. M. Bento Soares, bento-soares@uiowa.edu
                 The following repetitive elements were found in this cDNA
                 sequence: 32-117, >MER41B#LTR/MER4-group
                 Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES   Location/Qualifiers
source   1..686
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="UI-H-FT1-bib-1-22-0-UI"
                  /tissue_type="Aveolar Macrophage"
                  /dev_stage="Adult"
                  /lab_host="DH10B (Life Technologies)"
                  /clone_lib="NCI_CGAP_FT1"
                  /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                  modified polylinker; Site 1: EcoR I; Site 2: Not I;
                  NCI_CGAP_FT1 is a normalized cDNA library constructed from
                  a pool of 81 RNA samples from Alveolar Macrophages
                  challenged with different treatments. The library was
                  normalized according to Bonaldo, Lennon and Soares, Genome
                  Research, 6:791-806, 1996. First strand cDNA synthesis was
                  primed with an oligo-dT primer containing a Not I site.
                  Double stranded cDNA was ligated to an EcoR I adaptor,
                  digested with Not I, and cloned directionally into
                  pT7T3-Pac vector. The oligonucleotide used to prime the
                  synthesis of first-strand cDNA contains a library tag
                  sequence that is located between the Not I site and the
                  (dT)18 tail. The sequence tag for this library is
                  GGCCATGCCG. The tissue was provided by Dr. Gary W.
                  Hunninghake of the University of Iowa.
                  TAG_LIB=UI-H-FT1
                  TAG_TISSUE=Human Lung Aveolar Macrophage
                  TAG_SEQ=GGCCATGCCG"

```
RESULT 12
LOCUS       BI911134
BI911134
LOCUS       BI911134             672 bp    mRNA    linear    EST 16-OCT-2001
DEFINITION  603067546F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216770 5',
            mRNA sequence.
ACCESSION   BI911134
VERSION     BI911134.1  GI:16174696
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 672)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM11544 row: k  column: 11
            High quality sequence stop: 670.
FEATURES             Location/Qualifiers
     source          1..672
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5216770"
                     /tissue_type="leukocyte"
                     /clone_lib="NIH_MGC_118"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated adult donors. Library is oligo-dr primed
                     and directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
ORIGIN
```

BASE COUNT   143 a   180 c   174 g   188 t   1 others

Query Match
Best Local Similarity   99.9%; Score 671; DB 14; Length 686;
Matches 671; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy   4646 AGAACACGAGGCCCAGCCAGAGTACGGGAAGCGTTGTTCAGGTCAGCCAGTCCTGGGGACCGGA 4705
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 AGAACACGAGGCCCAGCCAGAGTACGGGAAGCGTTGTTCAGGTCAGCCAGTCCTGGGGACCGGA 627

Qy   4706 AGTGCAAAGATGGACTCTACCAAGCCCTGAAGAGAGCACCATCCTCACCTCATTTATGAAC 4765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    626 AGTGCAAAGATGGACTCTACCAAGCCCTGAAGAGAGCACCATCCTCACCTCATTTATGGAAC 567

Qy   4766 TCTGGGGAGGAGGCAGCAGCAAAAAAGGGACTCCTGCCACTCAGACTGAAGTATCAAACACCA 4825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 TCTGGGGAGGAGGCAGCAGCAAAAAAGGGACTCCTGCCACTCAGACTGAAGTATCAAACACCA 507

Qy   4826 GCCCTTGACCCTTGAGTCCTGGGCTGTTGGCTGCTCCTTCTTGTTTGGGTCTCAGTTTCTTTCTC 4885
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 GCCCTTGACCCTTGAGTCCTGGGCTGTTGGCTGCTCCTTCTTGTTTGGGTCTCAGTTTCTTTCTC 447

Qy   4886 TCCAAACAAAGTTGCCATCTGGTTGCTTCCACCACTAAAGTAATGAAACTTTTGATGATG 4945
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446 TGCAAACAAAGTTGCCATCTGGTTGCTTCCACCACTAAAGTAATGAAACTTTGATGATG 387

Qy   4946 CCTTTGCTGGGGCATTTATGTGTCCATGCCATGCCACAGGGGGCCCCAGTTCCAGGTGG 5005
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 CCTTTGCTGGGGCATTTATGTGTCCATGCCATGCCACAGGGGGCCCCAGTTCCAGGTGG 327

Qy   5006 CCTAACAGCATCTCAGGGAATGTCCATCTTCCATGCCAGGGATGCCACAGACCTCATAG 5065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 CCTAACAGCATCTCAGGGAATGTCCATCTTCCATGCCAGGGATGCCACAGACCTCATAG 267

Qy   5066 AGCCTCATCTGGTTGGGCCACAGCAGCCAAGCCTAGAGCCCTCCGATCCCATCCATCCAGGCCCA 5125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 AGCCTCATCTGGTTGGGCCACAGCAGCCAAGCCTAGAGCCCTCCGATCCCATCCAGGCCGCA 207

Qy   5126 AAGAGGAATAGGAGGGACATGGAACCATTTGCCTCTGGCTGTGTCACAGGGGTGAGCCCCA 5185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 AAGAGGAATAGGAGGGACATGGAACCATTTGCCTCTGGCTGTGTCACAGGGGTGAGCCCCA 147

Qy   5186 AAAATTGGGGTTCAGCGTGGGAGGGCCACGTGGATTCTTGGCTTTGTACAGGAAGAGATCTACA 5245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146 AAATTGGGGTTCAGCGTGGGAGGGCCACGTGGATTCTTGGCTTTGTACAGGAAGAGATCTACA 87

Qy   5246 AGAGCAAGCCACAGAGTAAAGTGGAGGAAGTTTATTCAGAAAATAAAGAGGGATATCACA 5305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     86 AGAGCAAGCCACAGAGTAAAGTGGAGGAAGTTTATTCAGAAAATAAAGAGGGATATCACA 27

Qy   5306 GCTCTTTTTAGAA 5317
          |||||||||||||
Db     26 GCTCTTTTTAGAA 15
```

BASE COUNT   138 a   187 c   189 g   158 t

Query Match
Best Local Similarity   99.1%; Score 633.6; DB 12; Length 672;
Matches 658; Conservative   0; Mismatches   4; Indels   2; Gaps   2;

```
Qy   3583 GCTTCCCANGTTGCTCAGGCTANTCTCAAAACTCTGGACGTGACGCAAGATCTTCCCAATT 3642
          |||||||| ||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      1 GCTTCCCANGTTGCTCAGGCTANTCTCAAAACTCTGGACGTGACGCAAGATCTTCCCAATG 60

Qy   3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGTCGGTGCCT 3702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGTCATGCCT 120

Qy   3703 TCTCCCT-GCCTCTCAAGGGGGACCTGCATACGGAAGGCCTTTGGGGACTGACGATGACTTCTG 3761
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCTCCTGGCCTCTCAAGGGGGACCTGCATACGGAAGGCCTTTGGGGACTGACGATGACTTCTG 180

Qy   3762 GGGCCCCACGGGGGCCTGTGGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTGTACCGAGT 3821
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGCCCCACGGGGGCCTGTGGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTGTACCGAGT 240

Qy   3822 TCACTTCCCTGTAGCTCAGGCGTCCTACCGCTGGGCCCAACACGGGTCGTCTGCTTCTTTGATGAG 3881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTTCCCTGTAGCTCAGGCGTCCTACCGCTGGGCCCAACACGGGTCGTCTGCTTTGATGAG 300

Qy   3882 AGAAGCGGTGACCGTTGAGATTGAATTCTGTGTGTGGGGACCAGTTCCTGGGTGAGATCAA 3941
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGAAGCGGTGACCCGTTGAGATTGAATTCTGTGTGTGGGGACCAGTTCCTGGGTGAGATCAA 360

Qy   3942 CCCACAGCACAGCTGGATGGTGGCAGGGCCCTCGTCTGCTGACACATCAAAGGCTGAGCCTGGAGC 4001
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCCACAGCACAGCTGGATGGTGGCAGGGCCCTCGTCTGCTGACACATCAAAGGCTGAGCCTGGAGC 420

Qy   4002 TGTGGAAGCTGTGCACCTCCCTCCACTTTGTGGCTCTCCAAGGGGGCGGCCATGTGGACACATC 4061
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGTGGAAGCTGTGCACCTCCCTCCACTTTGTGGCTCTCCAAGGGGGCGGCCATGTGGACACATC 480

Qy   4062 CCTGTTCCAAATGGCCCACTTTAAAGAGGAGGAGATGCTTCCTGGAGGAGAAGCCAGCCAGGGT 4121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCTGTTCCAAGCGCCCCACTTTAAAGAGGAGGAGATGCTTCCTGGAGGAGAAGCCAGCCA-GGT 539

Qy   4122 GGAGCTGCATCACATAGTTCTGAAAAACCCCAGTTCTCCCCCTTGGGGAGTCCTCCTGAA 4181
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCTGTTCCAAGCGCCCCACTTTAAAGAGGAGGAGATGCTTCCTGGAGGAGAAGCCAGCCA-GGT 539
```

```
Db    540 GGAGCTGCATCACATAGTTCTGGAAAACCCGAGCTTCTCCCCCTTGGGAGTCCTCCTGAA 599

Qy   4182 AATGGATCATAATGCCCTGGCTTCATTCCCGTTCGGTTGGTTGCTTACCACACG 4241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 AATGGATCATAATGCCCTGGCTTCATTCCCGTCACCTCGTTGGGTTGTTGCTTACCACACG 659

Qy   4242 CGTC 4245
          ||||
Db    660 CGTC 663
```

RESULT 13

Query Match   11.6%; Score 633; DB 12; Length 728;
Best Local Similarity   98.3%; Pred. No. 2.4e-75;
Matches 713; Conservative   0; Mismatches   5; Indels   7; Gaps   7;

```
Qy    885 CCATGAAATTCCGCGGCGGGGGGCTCAGAGAGAAGGGTTTTGAGACAGCTGCC 944
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 CCATGAAATTGCCGCGGGGGGTGCACCCA-GGCTCAGAGAGAAGGGTTTTGAGACAGCTGCC 64

Qy    945 TGACACATCTGGAGCGGCGCGTGGAGAGAAATTCTCTGCCTCACTCCTTACTCACAAGCTCTTCC 1004
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCCTCACTCCTCCTTACCCAAGCTCTTCC 124

Qy   1005 AAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAACGGCCCCCACCATCCACACAGC 1064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    125 AAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGCGCCCCACCATCCACACAGC 184

Qy   1065 AGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCGTACCCAGGATGCAGGAGGC 1124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 AGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCGTACCCAGGATGCAGGAGGC 244

Qy   1125 TCCTGGGACCCAAATGCCTCTGGATGAAAACGTCAGGAATTTACTACACAGAAATCAGGA 1184
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 TCCTGGGACCCAAATGCCTCTGGATGAAAACGTCAGGAATTTACTACACAGAAATCAGGA 304

Qy   1185 AAGAGAGAGAGAGAAAATCA-GAGAAAAGGAGCAGGGCCCCCATGGGCAGCGGGTTGGTGGGAACGC 1243
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 AAGAGAGAGAGAGAAATCACAGAGAAAAGGCAGGGCCCCCATGGGCAGCGGTGGGTTGGTGGGAACGC 364

Qy   1244 CCCCACAGGGGGCCACCAGCCTACAGCCCCACCACCCCACAGGGAGCCCTTCTGTGAGAG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 CCCCACAGGGGGCCACCAGCCTACAGCCCCACCACCCCACAGGGAGCCCTTCTGTGAGAG 424

Qy   1304 AGAGGCCTCTGTTCCACATGGCCCCTGAAAAAATGAGGATTTTAACAAAAAATTCACACAGC 1363
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 AGAGGCCTCTGTTCCACATGGCCCCTGAAAAAATGAGGATTTTAACAAAAAATTCACACAGC 484

Qy   1364 TGCTACTTCTTACAAAGACCTTCACCCCCCAGAAGACCAAGAATCCCCTGGTTCAAGGAGAAGCTGGC 1423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485 TGCTACTTCTTACAAAGATCTCACCCCAGAAGACCAAGATCCCCTGGTTCAAGGAGAAGCTGGC 544

Qy   1424 CTGA-TTATGTGGA-GGAGAATCGAGGACATTTAATTGAGATTCAGAGAAGGAGAAGCTGGCCC 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    545 CTGACTTATGTGGACGGAGAATCGAGGACATTTAATTGAGATTCAGAGAAGGACTTATTTGG-CC 603

Qy   1482 AGGCCTGGATACCCAAGAACCTTGCATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAA 1541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 AGGCCTGGGATACCCAAGAAACCTCGCATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAA 663

Qy   1542 GTCAACACTGGCCAGG--CAGCTGAAGGAAGCCTGGGGGAGAGGCCAGCTGTGTATGGGGACC 1600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 GTCAACACTGGCCAGGCCAGGTGAAGGAAGCCTGGGGGAGAGG-CAGCTGTATGGGGACA 722

Qy   1601 GCTGTC 1605
          ||||||
Db    723 GTTCC 727
```

RESULT 14

source 1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/clone_lib="Chondrosarcoma"
/db_xref="taxon:9606"
/clone="IMAGE:5845056"
/tissue_type="Adult"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Ell"
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
adaptor; Site_1: EcoR I; Site_2: Not I;
NCI CGAP Ell is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTTGCAC."
TAG_LIB=UI-H-Ell
TAG_TISSUE=chondrosarcoma
TAG_SEQ=ACACTTGCAC

BASE COUNT 143 a 170 c 164 g 178 t
ORIGIN

Qy Query Match 11.5%; Score 628.4; DB 12; Length 655;
 Best Local Similarity 99.7%; Pred. No. 1e-74;
 Matches 640; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 4676 TGTTCAGCTTGAGCCAGTCCTGGGACCCGGAAGTGCAAAGATGGACTCTACCAAGCCCTGA 4735
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 655 TGTTCAGCTTGAGCCAGTCCTGGGACCCGGAAGTGCAAAGATGGACTCTACCAAGCCCTGA 596

Qy 4736 AGGAGACCCATCCTCACCTCATTATGGAACTCTGGGAGGAAGGGCAGCAAAAAGGGACTCC 4795
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 595 AGGAGACCCATCCTCACCTCATTATGGAACTCTGGGAGGAAGGGCAGCAAAAAGGGACTCC 536

Qy 4796 TGCCACTCAGCGTGAAGTATCAAACACCAGCCCTTGACCCTTGAGTCTGGCTTTGGGCT 4855
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 535 TGCCACTCAGCGTGAAGTATCAAACACCAGCCCTTGACCCTTGAGTCTGGCTTTGGGCT 476

Qy 4856 GACCCTTCTTGGGTCTCAGTTTCTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTC 4915
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 475 GA--CCTTCTTGGGTCTCAGTTTCTTCTCTGCAAACAAGTTGCCATCTGGTTTGCCTTC 417

Qy 4916 CAGGACTAAAGTAATGGAACTTTGATGAATGCCTTGCGTGGGCATTATGTGTCCATGCCAG 4975
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 416 CAGGACTAAAGTAATGGAACTTTGATGAATGCCTTGCGTGGGCATTATGTGTCCATGCCAG 357

Qy 4976 GGATGCCACAGAGGGGGGCCCCAGTCTCCAGGGGTGGCCTAACAGACATCTCAGGGAATGCCCATCTG 5035
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 356 GGATGCCACAGAGGGGGGCCCCAGTCTCCAGGGGTGGCCTAACAGACATCTCAGGGAATGCCCATCTG 297

Qy 5036 GAGCTGGCAAGACCCCTGCGAGACCTCATAGAGCTTCATCTGGTTGGTGGCCACAGCCAAGC 5095
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296 GAGCTGGCAAGACCCCTGCGAGACCTCATAGAGCTTCATCTGGTTGGTGGCCACAGCCAAGC 237

Qy 5096 CTAGAGCCCTCCGGATCCCATCCAGGCGCAAAGAGAGGAATAGGAGGACATGGAACCATTT 5155
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 236 CTAGAGCCCTCCGGATCCCATCCAGGCGCAAAGAGAGGAATAGGAGGACATGGAACCATTT 177

Qy 5156 GCCTTCTGGCTGTGTGTCACAGGGTGAGCCCCCAAAAATTGGGGTTCACGTGGGAGGGCCACGTG 5215
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 176 GCCTTCTGGCTGTGTGTCACAGGGTGAGCCCCCAAAAATTGGGGTTCACGTGGGAGGGCCACGTG 117

Qy 5216 GATTCTTGGCTTTGTACAGGAGATCTACAAGAGCAAGCCAACAGAGTAAAGTGGAAGGA 5275
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 116 GATTCTTGGCTTTGTACAGGAGATCTACAAGAGCAAGCCAACAGAGTAAAGTGGAAGGA 57

Qy 5276 AGTTTATTCAGAAAAATAAAGGAGTATCAGGCTCTTTAGAA 5317
     |||||||||||||||||||||||||||||||||||||||||
Db 56 AGTTTATTCAGAAAAATAAAGGAGTATCAGGCTCTTTAGAA 15

LOCUS BM152462 655 bp mRNA linear EST 30-NOV-2001
DEFINITION TCBAP1B9080 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP9080, mRNA
 sequence.
ACCESSION BM152462
VERSION BM152462.1 GI:17176729
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 655)
AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R.Jr.;
 Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
TITLE Pediatric Leukemia cDNA Sequencing Project (2001)
JOURNAL Unpublished
COMMENT Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@txccc.org
 Seq primer: M13 primer

FEATURES Location/Qualifiers
source 1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP9080"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"
/note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGGCGGGCGGCGGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGAGCTCGGATCCCGCGCGCGCGCAATAATAATAT(C)3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasakin, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 177 a 180 c 180 g 116 t 2 others
ORIGIN

Qy Query Match 11.5%; Score 623.6; DB 12; Length 655;
 Best Local Similarity 98.8%; Pred. No. 4.5e-74;
 Matches 648; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 1087 CCACCTCAGCCCAGCCTAGCCACCCAGAGAGGACCAGAGGAGGCTCCTGGGACCCAGATGGCCTCTG 1146
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CCACCTCAGCCCAGCCTAGCCACCCAGAGAGGACCAGAGGAGGCTCCTGGGACCCAGATGGCCTCTG 60

Qy 1147 GATGAAACGTCAGGAAATTTACTACACAGAGAAATCAGAGAAAGAGAGAGAGAAAATCAGAG 1206
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GATGAAACGTCAGGAAATTTACTACACAGAGAAATCAGAGAAAGAGAGAGAGAAAATCAGAG 120

```
Qy   1207  AAAGGCAGGCCCCCCATGGGCCAGCGGTGGTGGTAGGAAGCCCCCAGGGCGCACACCAGGCCTTA  1266
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ATTGGCAGGGCCCCATGGGCCAGGCGGTGGTAGGAACGCCCCACACCAGGCCTTA            180

Qy   1267  .CAGCCCCACCACCACCACCCATGGGAGCCTTCTGTGTGAGAGAGAGCCTCTGTTCCACATGGCCC  1326
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CAGCCCCACCACCACCACCCATGGGAGAGCCTTCTGTGTGAGAGAGAGCCTCTGTTCCACATGGCCC  240

Qy   1327  TGGAAAAATGAGGATTTTAACCAAAAAATTCACACAGCCTGCTACTTCCTACAAAGACCTCAC  1386
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  TGGAAAAATGAGGATTTTTAACCAAAAAATTCACACAGCCTGCTACTTCCTACAAAGACCTCAC   300

Qy   1387  CCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGGCCTGATTAATGTGGAGGAGAATCGA  1446
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGGCCTGATTAATGTGGAGGAGAATCGA   360

Qy   1447  GGACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCTGGATACCCAAGAACCTGC  1506
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GGACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCTGGATACCCAAGAACCTGC   420

Qy   1507  ATAGTCATACTGCACGGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCCAGGAAGGTGAAG  1566
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  ATAGTCATACTGCACGGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCCAGGAAGGTGAGG   480

Qy   1567  GAAGCCTGGGGGGAGGAGGCGCAGCTGTTATGGGGGACCGCTTCCAGCATGTCTTCTACTTCAGC  1626
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  NAAGCCTGNGGGAGGAGGCCAGCTGTTATGGGGGACCGCTTCCAGCAATGTCTTCTACTTCAGC   540

Qy   1627  TGCAGAGAGCTGGCCCAGTCCAAGTGTGGTGAGTCTCGCTGAGCTCATCGGAAAAAGATGGG  1686
                 |||||||||||||||||||||||||||||||||||||||||||
Db    541  TGCAGAGAGCTGGCCCAGTCCAAGGGTGTGAGTCTCGCTGAGCTCATCGG-AAAGATGGG   599

Qy   1687  ACAGCCCACTCCGGCTCCCATTAGACAGATCCTGTCTAGGCCAGAGACGGGCTGCTCTT  1742
                 ||||||||||||||||||||||||||||||||||||||||||||||||
Db    600  ACAGCCACTCCGGCT-CCATTAGACAGATCCTGTCTAGCCCAGAGCGGGCTGCTCTT   654
```

Search completed: January 29, 2004, 13:05:49
Job time : 10855.3 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 29, 2004, 03:02:31 ; Search time 291.391 Seconds (without alignments)
8246.264 Million cell updates/sec

Title:
Perfect score: 5444
Sequence: 1 gcccagggcctggagagt.........aggaataagaagttacctac 5444
US-09-996-617-1

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	165.8	3.0	1371	2	US-08-910-731-1	Sequence 1, Appli
2	165.8	3.0	1371	2	US-08-795-395-1	Sequence 6, Appli
3	163.6	3.0	1371	2	US-08-910-731-7	Sequence 14, Appl
4	163.6	3.0	1386	2	US-08-910-731-5	Sequence 1, Appl
5	156.2	2.9	1374	2	US-08-910-731-3	Sequence 5, Appli
6	156.2	2.9	1374	2	US-08-795-395-3	Sequence 3, Appli
7	115	2.1	585	4	US-09-340-620A-50	Sequence 50, Appl
8	115	2.1	740	4	US-09-340-620A-48	Sequence 48, Appl
9	106.2	2.0	579	4	US-09-340-620A-62	Sequence 62, Appl
10	106.2	2.0	777	4	US-09-340-620A-60	Sequence 60, Appl
11	62	1.1	7218	1	US-08-232-463-14	Sequence 14, Appl
12	59.4	1.1	257	4	US-09-016-434-208	Sequence 208, App
13	57.2	1.1	3116	4	US-09-904-615-43	Sequence 43, Appl
14	49	0.9	337	1	US-08-594-031-79	Sequence 79, Appl
15	48.6	0.9	1878	4	US-09-465-558-39	Sequence 39, Appl
16	48	0.9	1582	3	US-08-545-196B-10	Sequence 10, Appl
17	48	0.9	1582	3	US-08-545-196B-12	Sequence 12, Appl
18	47.2	0.9	17949	3	US-09-087-465-3	Sequence 3, Appli
19	46.2	0.8	599	3	US-09-328-111-147	Sequence 147, App
20	46.2	0.8	1441	3	US-08-821-994-63	Sequence 63, Appl
21	46.2	0.8	4141	4	US-09-245-281-42	Sequence 42, Appl
22	46.2	0.8	4141	4	US-09-207-359B-42	Sequence 42, Appl
23	46.2	0.8	4141	4	US-09-340-620A-42	Sequence 42, Appl
24	46.2	0.8	924	1	US-08-468-709B-1	Sequence 1, Appli
25	45.8	0.8	924	2	PCT-US93-03936-1	Sequence 1, Appli
26	45.8	0.8	924	5	PCT-US93-03936-1	Sequence 1, Appl
27	45.8	0.8	1717	1	US-08-468-709B-6	Sequence 6, Appli
28	45.8	0.8	1717	2	US-08-241-664B-6	Sequence 6, Appli
29	45.8	0.8	1717	5	PCT-US93-03936-6	Sequence 6, Appli
30	45.6	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
31	45.2	0.8	289	1	US-09-007-005-17	Sequence 17, Appl
32	45.2	0.8	289	3	US-09-244-796-17	Sequence 17, Appl
33	45	0.8	1508	4	US-09-039-046-1	Sequence 1, Appli
34	45	0.8	111282	4	US-09-754-250-3	Sequence 3, Appli
35	44.4	0.8	1190	4	US-09-390-207-1	Sequence 1, Appli
36	44.4	0.8	3715	4	US-08-234-245-1	Sequence 1, Appli
37	43.8	0.8	467	2	US-08-841-349-18	Sequence 18, Appl
38	43.8	0.8	1868	1	US-08-658-883B-1	Sequence 1, Appli
39	43.8	0.8	1868	3	US-08-676-610B-26	Sequence 26, Appl
40	43.8	0.8	3437	4	US-09-489-847-89	Sequence 89, Appl
41	43.2	0.8	1342	4	US-08-860-339-17	Sequence 17, Appl
42	43.2	0.8	1534	1	US-09-300-903A-6	Sequence 6, Appli
43	43.2	0.8	1534	4	US-08-988-197-6	Sequence 6, Appli
44	42.8	0.8	1143	4	US-09-220-132-166	Sequence 166, App
45	42.8	0.8	1700	2	US-08-897-340-4	Sequence 4, Appli

RESULT 1
US-08-910-731-1
Sequence 1, Application US/08910731
Patent No. 5932440

GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 2005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS

ALIGNMENTS

; LOCATION: 1..1368
US-08-910-731-1

Query Match   3.0%; Score 165.8; DB 2; Length 1371;
Best Local Similarity   56.0%; Pred. No. 8.7e-34;
Matches 314; Conservative   0; Mismatches 247; Indels   0; Gaps   0;

```
Qy   2920 CTCTTCTCCGTCTCTCAAGGTCACCAGAAAACCTGAAGGAGCTAAGTGGAAACTCG 2979
          ||||| ||  ||||| |||   ||| || ||  || || |||||  ||||||| ||
Db    811 CTCTGCCCGTCTCCTCCAGGCCAAGGAGACCTCAGTCTGGCGGGCAACAAG 870

Qy   2980 CTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACAGGCCCCTCGCTCTCCTG 3039
          ||| || |||||  || ||||  | |   ||||||||||||| ||  ||||| || || ||
Db    871 CTGGGCGACGAGGCGCCCGGCTTGCGTGTCCAGGAGCGGCCCCGGCTTGCCCAGCTG 930

Qy   3040 GAGACCCTGCGGTTGGCTGGCTGTGGCCCTCACAGCTGAGGACTGCCAAGGACCTTGCCTTT 3099
          | | |||  |||||   ||| | || || ||| ||| ||| |||||| |||||||| |||
Db    931 GAGTCCCTGTGGGTGAAGTCCTCGCAGCCTCACGGCGGCCCCGGCTTGCCCAGCTTG 990

Qy   3100 GGGCTGAGGAGGCCAAACACACAGCTCACGGACTGAGGACTGGAGTTCAATGTGCTCACGGAT 3159
          |||  |||||   |||  ||| |  |  || ||| | || || || |   ||||| || ||
Db    991 ATGCTGACCCAGGAACAACAGCATCTCCTGAACTTCAGTTGAGCAGCACAAGCTGGGGTAC 1050

Qy   3160 GCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCGGCCGAGCTGCAAGCTACAGCGACTG 3219
          | || || ||| |||| | ||| |||||||  ||||||||| |||| ||| | |||||| ||
Db   1051 TCTGGCATCCAGGAGCTGTCGCCAGGCCCTGAGCCACCACTGCCAGCGCTGCTC 1110

Qy   3220 CAGCTGGTGCAGCTGTGGGCCTCACGTCTGACTGCTGCCAGGAACCTGGCCTTCTTGTGCTTAGT 3279
          ||||||  ||||| |  ||||||||| |||| || |||||| || |||||  ||||| |||
Db   1111 TGTCTTGGGGACTGTGAGGTGACCAAACGCGGTGTCGCAGCGGTGCTGCCCTTCGCTTCCTGCTG 1170

Qy   3280 GCCAGCCCCAGCCTGAAGGAGCTAGACGCTGCAGAACAACACTGGATGACGTTGGCGTG 3339
          ||||| || |||| || ||||| || |||||| || | |||| | ||| ||| |||||
Db   1171 GCCAACCGGCAGCCTGCGAGACTGACCTGAGCAACTGTGTGGGCGACGCCCGGCCGTC 1230

Qy   3340 CGACTGCTTCTGTGAGGGGCTCAGGCATCCTGCCTGCAAAACTCATACGGCCTGGGGGCTGGAC 3399
          |||| | | | |||| ||||| |||  || | || || ||||| | ||||||| ||||||
Db   1231 CTGCAGCTGCTGGGGAGCCTGACCCGGGCCCGGCTGAGGCGGCTGTGCCTGGGCTGGTCTTGCTAC 1290

Qy   3400 CAGACAACTCTGAGTGATCAGAGATGAGGGCAGGAACTGAGGGACCTGGACAGAGAAACT 3459
          || | ||||||||| | || |||  ||| |||||| ||||| |||||| ||| ||||
Db   1291 GACACCTACTGGACGAGGAGGAGTGGAGAGGACCCGCCTGCGAGGCCCCTGGAGGGGAGCAAGCCC 1350

Qy   3460 CAGCTGCTTCATCTTCAGCAGA 3480
          ||||  ||| |||| |||||||
Db   1351 GGCCTGAGGGTCATCTCCTGA 1371
```

; RESULT 2
US-08-795-395-1
; Sequence 1, Application US/0795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
US-08-795-395-1

Query Match   3.0%; Score 165.8; DB 2; Length 1371;
Best Local Similarity   56.0%; Pred. No. 8.7e-34;
Matches 314; Conservative   0; Mismatches 247; Indels   0; Gaps   0;

```
Qy   2920 CTCTTCTCCGTCTCTCAAGGTCACCAGAAAACCTGAAGGAGGCTGAACTAAGTGGAAACTCG 2979
          ||||| ||  |||||  ||    | || || || || |||||| || |||||  ||||| ||
Db    811 CTCTGCCCGTGTCCTCCAGGCCAAGGAGACCTCAGTCTGGCGGGCAACAAG 870

Qy   2980 CTGAGCGACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACAGGCCCCTCGCTCTCCTG 3039
          ||| || |||||  || ||||  | |   ||||||||||||| ||  |||| ||  ||  ||
Db    871 CTGGGCGACGAGGCGCCCGGCTTGCGTGTCCAGGAGCGGCCCCGGCTTGCCCAGCTG 930

Qy   3040 GAGACCCTGCGGTTGGCTGGCTGTGGCCCTCACAGCTGAGGACTGCCAAGGACCTTGCCTTT 3099
          | | |||  |||||   ||| | || || ||| ||| ||| |||||| |||||||| |||
Db    931 GAGTCCCTGTGGGTGAAGTCCTCGCAGCCTCACGGCGGCCCCGGCTTGCCCAGCTTG 990

Qy   3100 GGGCTGGGAGGCCAAACACCTTTGCCAGAGACTGAGACAGCGGCCGAGCTGCAAGCTACAGCGACTG 3159
          |||   ||||  ||| ||| | ||| |||||| |  ||||||||| |||| ||| || |||||| ||
Db    991 ATGCTGACCCAGGAACAACAGCATCTCCTGAACTTCAGTTGAGCAGCACAAGCTGGGGTGC 1050

Qy   3160 CAGCTGGTCAGCTGTGGGCCTCACGTCTGACTGCTGCCAGGAACCTGGCCCTCTGTGCTTAGT 3219
          ||||||  ||||| |  ||||||| | ||| || |||||| || |||||  ||| || |||
Db   1051 TCTGGCATCCAGGAGCTGTCGCCAGGCCCTGAGCCACCACTGCCAGCGCTGCTC 1110

Qy   3220 CAGCTGGTGCAGCTGTGGGCCTCACGTCTGACTGCTGCCAGGAACCTGGCCTTCTTGTGCTTAGT 3279
          ||||||  ||||| |  ||||||||| |||| || |||||| || ||||| ||||| |||
Db   1111 TGTCTTGGGGACTGTGAGGTGACCAAACGCGGTGTCGCAGCGGTGCTGCCCTTCGCTTCCTGCTG 1170

Qy   3280 GCCAGCCCCAGCCTGGAGGAGCTAGACCTGCTGGGGATGAGCAACAACTGTGTGGGCGACGCC 3339
          ||||| ||| |||| |||||| || |||||||||| |||| |||||| ||||| |||||| ||
Db   1171 GCCAACCGGCAGCCTGCGAGACTGACCTGAGCAACTGTGTGGGCGACGCCCGGCCGTC 1230

Qy   3340 CGACTGCTTCTGTGAGGGGCTCAGGCATCCTGCCTGCAAAACTCATACGGCCTGGGGGCTGGAC 3399
          |||| | |||| |||| ||||| |||  || | || || ||||| | ||||||| ||||||
Db   1231 CTGCAGCTGCTGGGGAGCCTGACCCGGGCCCGGCTGAGGCGGCTGTGCCTGGGCTGGTCTTGCTAC 1290

Qy   3400 CAGACAACTCTGAGTGATCAGAGATGAGGGCAGGAACTGAGGGACCTGGACAGAGAAACT 3459
          || | ||||||||| | || |||  ||| |||||| ||||| |||||| ||| ||||
Db   1291 GACACCTACTGGACGAGGAGGAGTGGAGAGGACCCGCCTGCGAGGCCCCTGGAGGGGAGCAAGCCC 1350

Qy   3460 CAGCTGCTCATCTTCAGCAGA 3480
          ||||  ||| |||| |||||||
Db   1351 GGCCTGAGGGTCATCTCCTGA 1371
```

; RESULT 3
US-08-910-731-7
; Sequence 7, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:

Left column

```
;  APPLICANT: CHATTERJEE, DEB K.
;  APPLICANT: SHANDILYA, HARINI
;  TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
;  NUMBER OF SEQUENCES: 16
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;     STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;     CITY: WASHINGTON
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3934
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/910,731
;     FILING DATE: (Herewith)
;     CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/795,395
;     FILING DATE: 04-FEB-1997
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/794,546
;     FILING DATE: 03-FEB-1997
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/024,057
;     FILING DATE: 16-AUG-1996
;  ATTORNEY/AGENT INFORMATION:
;     NAME: ESMOND, ROBERT W.
;     REGISTRATION NUMBER: 32,893
;     REFERENCE/DOCKET NUMBER: 0942.3440003
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-371-2600
;     TELEFAX: 202-371-2540
;  INFORMATION FOR SEQ ID NO: 7:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 1371 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: both
;        TOPOLOGY: both
;        MOLECULE TYPE: cDNA
US-08-910-731-7

Query Match           3.0%;  Score 163.6;  DB 2;  Length 1371;
Best Local Similarity 55.7%;  Pred. No. 3.3e-33;
Matches 313;  Conservative 0;  Mismatches 249;  Indels 0;  Gaps 0;

Qy 2919 TCICTTCTCCGTCCTCAAGGTCACCAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACTC 2978
Db  810 TCTGTGCCCTGTCCTCAGGGCCAAGGAGGAGCCTGAAGGAGCTCAGCCTGGCCGGCAACGA 869

Qy 2979 GCTGAGCCACTCTCGCAGTGAAAGAGTCTTTGTAAGACACCTGAGAGAGGCCCCTCGCTCC 3038
Db  870 GCTCGGGGATGAGGGTGCCCGACTGCTGTGTGAGAACCTTGAGGCTGCCAGCT 929

Qy 3039 GGAGACCCCTGGCGTTGGCTGGCTCGTGGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCTT 3098
Db  930 GGAGTCGCTGTGGGTGAAGTCCTCCAGCTTCACAGCCGCTGCTGCCGAGAACCTGCAGCTC 989

Qy 3099 TGGGCTGAGAGCCAACCAGACCCTGACCGGACCTGGACTTGCAATGTGCTCACGGA 3158
Db  990 AGTGCTGGCGCCCAGAACACAGGTTTCTTCCTGCAGGACCTACAGCGACT 1049

Qy 3159 TGCTGGAGCCAAAACACCTTTGCCAGAGACTGAGACAGCCGAGACTGAAGACCTGGACGACT 3218
Db 1050 TGCGGGGCGTGCGGGAGCTGTGCCCAGGCCCTGGCCTTCTGTTGTCTTAG 1109

Qy 3219 GCACCTGGTTCAGTCGTGCTGCCTCACGTCTGCTGCCCAGGACCTTGGCCTTCTGTTGTCTTAG 3278
Db 1110 CTGGTTGGGCCGACGTGAGTGACAGCAGCTTGCAGCAGCGCTTCGCCGCCGCAACCCTGTT 1169
```

Right column

```
RESULT 4
US-08-910-731-5
Sequence 5, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:
;  APPLICANT: CHATTERJEE, DEB K.
;  APPLICANT: SHANDILYA, HARINI
;  TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
;  NUMBER OF SEQUENCES: 16
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;     STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;     CITY: WASHINGTON
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3934
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/910,731
;     FILING DATE: (Herewith)
;     CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/795,395
;     FILING DATE: 04-FEB-1997
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/794,546
;     FILING DATE: 03-FEB-1997
;  PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/024,057
;     FILING DATE: 16-AUG-1996
;  ATTORNEY/AGENT INFORMATION:
;     NAME: ESMOND, ROBERT W.
;     REGISTRATION NUMBER: 32,893
;     REFERENCE/DOCKET NUMBER: 0942.3440003
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-371-2600
;     TELEFAX: 202-371-2540
;  INFORMATION FOR SEQ ID NO: 5:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 1386 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: both
;        TOPOLOGY: both
;        MOLECULE TYPE: cDNA
US-08-910-731-5

Query Match           3.0%;  Score 163.6;  DB 2;  Length 1386;
Best Local Similarity 55.7%;  Pred. No. 3.4e-33;
Matches 313;  Conservative 0;  Mismatches 249;  Indels 0;  Gaps 0;

Qy 2919 TCICTTCTCCGTCCTCAAGGTCACCAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACTC 2978

Qy 3279 TGCCAGCCCCAGCCTGAAGAGGAGCTAGACTCGACAGCAGAGAACAAACCTGGATGACGTTGGCGT 3338
Db 1170 GGCCAACCACAGCCCTGCGTGAGCCTGGACCCCAGCAACAACTGCCTGGGGGGACGGCGCAT 1229

Qy 3339 GCGGACTGGCTCTGTGAGGGGGCTGAGGGCATCCTGCCTCGCCCAAACTCATACGCCCTGGGGCTGGA 3398
Db 1230 CCTGCAGCTGGTGGAGAGCGGTCCGGCCAGCCCAGCCTGCCTCCTGGACAGCAGCTGGTCCTGTA 1289

Qy 3399 CCAGACAACTCTGAGTGATGATAGGCAGGCAGGAAGTGAGGGCCCTGGAGCAGGAGAAACC 3458
Db 1290 CGACATTTACTGGTCTGAGGAGGATGGAGGACCCGGCTGCAGGCGCCCTGGAGAAGGACAAGCC 1349

Qy 3459 TCAGCTGCTCATCTTCAGCAGA 3480
Db 1350 ATCCCTGAGGGTCATCTCCTGA 1371
```

Qy   825 TCTGTGCCGTGTCCTCAGGGCCGAAGGAGGAGCCTGAAGGAGTCAGCCTTGGCCGCGCAACGA 884
Db 2979 GCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCTGAGACGCCCTTCGCTTGCCCTCCT 3038
Db   885 GCTGGGGGATGAGGGTGCCCGACTGCTGTGTGAGACCCTTGCGAACCTGGCCTGCCAGCT 944
Qy 3039 GGAGACCCTGCCGGTTGGCTGGCTGTGGGCCTTCACAGACTGCAAGGACCTTGCCTT 3098
Db   945 GGAGTCGCTGTGGGTGAAGTCCCAGCTTCACAGCCGCCGCCACTTCAGCTC 1004
Qy 3099 TGGGCTGAGAGCCAACCAGACCGAGCTGGACCTGGACTTCAATGTGCTCACGGA 3158
Db 1005 AGTCCTGGCCCAGAACAGGTTTCTCCTGACGAGCTACAGAGATAAGCAACACAGGCTGGAGGA 1064
Qy 3159 TGCTGGGAGCCAAAACACCTTTGCCAGAGACTGAGACGAGACAGCGACT 3218
Db 1065 TGCGGGGCGTGCCGGGAGCTGTGCCAGGGCCTGGCCTCTGTGCTGCGGGGTGCT 1124
Qy 3219 GCAGCTGGTCAGCTGTGTGGCCTCACGTCTGACTGCTGCCAGGACCTGGCCTCTGTGGCTTAG 3278
Db 1125 CTGGTTGGCGACTGCGATGTGACGAGCAGCAGCGTGCACCCTGTT 1184
Qy 3279 TGCCAGCGCCCAGCCTGAAGGAGCTTAGACCTGCAGCAGAACAACCTGGATGACGTTGGCGT 3338
Db 1185 GGCCAACCACAGCCTGCGTGAGCTGGACCTCAGACAACAACTGCCTGGGGACCGCGGGCAT 1244
Qy 3339 GCGACTGCTCTGTGAGGGGCTCAGGGACATCCGTGCCGCAAAACTCATACGGCCTTGGGGGCTGGA 3398
Db 1245 CCTGCAGCTGGTGGAGGAGGCCTCCGGCACCTTCACGCCGCCCTGCTGCCCCCACTTCAGCTC 1304
Qy 3399 CCAGACAACTCTCGAGTGATGAGTGAGGGCCAGGAACTGAGGGCGCAGGAGAAACC 3458
Db 1305 CGACATTTACTGGTCTGAGGAGATGGAGGACCCGCGCTGCAGGGCCCTGCAGGGCCTGGGAGGCAGGACAAGCC 1364
Qy 3459 TCAGCTGCTCATCTTCAGCAGA 3480
Db 1365 ATCCCTGAGGGTCATCTCCTGA 1386

Query Match   2.9%;   Score 156.2;   DB 2;   Length 1374;
Best Local Similarity 54.2%;   Pred. No. 3.1e-31;
Matches 317;   Conservative 0;   Mismatches 268;   Indels 0;   Gaps 0;

Qy 2896 GTCACAGATCCCTATTGGGCAGATTCTCTTCTCCGTCCTCAAGGTCACCAGAAAACTTGAAG 2955
Db   787 GTCACTGCAGAAGGCTGCAAGGACACCTGTGCCGTGTCCTCAGACGCCAAGCAGAGACCTGAAG 846
Qy 2956 GAGCTGGACCTAAGTGGGAAAACTCGCTGAGGACCACTCTGCAGTGAAGAGTCTTTGTAAGACC 3015
Db   847 GAACTCAGCCTAGCTGGGCAAATGAGCTGAAGGATGAGGGGTGCCCAACTGCGTGTGAGAGC 906
Qy 3016 CTGAGACGCCCTCGCTGCTCCTGGAGAGACCCGGTTGGCTGGGCTGGGCCTCACGGCT 3075
Db   907 CTGTTAGAGCCTGGCTGTGTCAGCTCGAGTCACTGTGGTAAAGAACCTTGTGAGCTTCACAGCT 966
Qy 3076 GAGGACTGCAAGGAGGACCTTGCCTTTGGGCGTGCTGAGAGCCAACCAGACCCTGGAC 3135
Db   967 GCCTCTTCGTCCCCACTTCTTCGCTCGGTGTTGACCAAAAACAGTTCTTCGTTTTGGAGTTGCAA 1026

Qy 3136 CTGAGCTTCAATGTGCTCACGGATGCCTGGAGACCAAAACACCTTTGCCAGAGACTGAGCAG 3195
Db 1027 ATGAGCAGCAACCCGCCTGGGAGAGACTCGGGAGGTCGTGCAAGGGCCCTGGCGTAT 1086
Qy 3196 CCGAGCTGCAAGCTACAGCGGACTGCCAGCGTGGTCAGCGTGGGCCTCAGTGCTGCCTGC 3255
Db 1087 CCGGACACACAGTGCTGCGCGTGTGCTGCTTTGGGCGTGGGAGGGCTGCTGGGAC 1146
Qy 3256 CAGGACCCTGGCCTCTTGTGTTGCTTAGTGCCAGCCCCAGCCTGAAGGAGCTGAGCAG 3315
Db 1147 AGCAGCCTTGCCACTGTCTCCTGCGACTGCCAACCGCAGCTTGAGGGGAAACTGGGAC 1206
Qy 3316 AACAACCTGGATGACGTTGGCGTGCGACTGCTCTGTGAGGGGCTCAGGGCATCCTGCCTGC 3375
Db 1207 AACTGCATGGGGGACAAACGTGTCCTACAACTGTGTGGGAGAGGCCCTCAAAAACAGCCCAGCTGC 1266
Qy 3376 AAAACTCATACGGCCTTGGGGGCTGGACCAGACAACTCTGAGTGATGAGTGAGGGCCAGGAACTG 3435
Db 1267 ATCCTTCAGCAGCTTGTCCTGTATGACATTTACTGGACGGATGAGGTGGAAGACCAGCTT 1326
Qy 3436 AGGGGCCCTGGACGAGGAGGAAACCTCAGCTGCCTGCCTGCCTCATCTTCAGCAGA 3480
Db 1327 CGGGGCCCTGGAGGGAGGAGGAAAGGCCATCCCTGAGGATCATTTCCTGA 1371

; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368

US-08-795-395-3

Query Match   2.9%;   Score 156.2;   DB 2;   Length 1374;
Best Local Similarity 54.2%; Pred. No. 3.1e-31;
Matches 317;   Conservative 0;   Mismatches 268;   Indels 0;   Gaps 0;

Qy 2896 GTCACAGAATGCCTAATTGGGACAGATTCTTCTTCCGGTCCTCAAGGTTCACCAGAAAACTGGAAG 2955
          |||||| |||||| |||| |||| ||| | | |||| ||||||||||||| |||||||||
Db   787 GTCACTGCAGAAGGGCTGCAAGGAACCTGTGCCGTGTTCTCAGAGCCAAGCCAAGCCTGAAG 846

Qy 2956 GAGCTGGACCTTAAGTGGGAAAACTCGCTGAGCGACCACTCTGAGTAGAAGAGTCTTTAAGACC 3015
          |||| ||||||| |||||| |||| |||| ||||| ||| || |||| |||| |||||||
Db   847 GAACTCAGCCTAGCTGGCAATGAGCTGAAGGATGAGGGTGCCCAACTGCTGTGTGTGTAAGAGC 906

Qy 3016 CTGAGAACGCCCTCGCTGTGCCTCCTGGAGAGCCCTGGCTGTTGGCTGTGGCCTCACAGCT 3075
          ||| ||||||||||||| || ||||||| |||||||||||||||||||||||||||
Db   907 CTGTTAGAGGCTGGCTGCTCAGCTGGAGTCACTGTGGGTAAAAGACCTTAGCCTCACAGCT 966

Qy 3076 GAGGACTGCAAGGACCTTGCCCTTTGGGCTGAGAGCCAACCAGACCCCTGACCGCCTGGGAC 3135
          ||||||| ||||||| ||| ||||||| ||| ||||||| |||||||||||| || |||
Db   967 CTGTTAGAGGCTGCGTCTGGCTGGAGTCACTGTGGGTAAAAGACCTTAGCCTCACAGCT 1026

Qy 3136 CTGAGCTTCAAATGTGCTCACGGATGCTGGGACCAAAACACTTTGCCAGAGACTGAGACAG 3195
          ||| |||||||||||||||||| |||||||||||| ||| |||||| ||| ||||||
Db 1027 ATGAGCAGACAACCCGGCTGGGAGACTCGGGAGTCGTGGAGCTTTGCAAAGGCCCTGGGCTAT 1086

Qy 3196 CCGAGCTGCAAGCTACAGCGACTGCAGCTGGTCAGCTGTGGCCTCAGTGTTGACTGCTGC 3255
          ||| || |||||||||||||||||| ||||||||| ||||||||| |||||||||||||
Db 1087 CCGGACACAGTGCTCCGGTGTCGTCCTTTGGCGTGATGTAGACACAGAGAGGGGCTGC 1146

Qy 3256 CAGGACCTGGCCTTCTTGTTGCTTAGTGCCAGCCCCAGCCTGAAGGAGTCTAGACCTTCAGCAG 3315
          ||||||||||||| ||||||||| ||||| |||| |||| |||||| ||| |||||||||
Db 1147 AGCAGCCTTCCCACTGTCTCCTGCTGCCGCAGCTTGAGGGGAACTGGACCTCAGTAAC 1206

Qy 3316 AACAACCTGGATGAGACGTTGGCGTGCCGACTGCTCTTGTGAGGGGGCTCAGGCGATCCTGCCTGC 3375
          |||| |||| ||| |||| |||||| |||| | ||||||||| ||||| |||||||| |||||||
Db 1316 CGGGGCCCTGGGAGCAGGAGAAAGGCCATCCCTGAGGATCATTCCTGA 1371

Db 1207 AACTGCATGGGGGGGCACAACGGTGTCCTACAACTGCTGCTGGAGAGCCTCAAACAGCCCAGCTGC 1266
          | ||||| |||||||| ||||||| ||| | | |||||| || |||| ||| ||| ||||||
Qy 3376 AAACTCATACGCCCTGGGGCTTGGACCAGACAACTCTCGAGTGATGAGGATGAGAGGCAGGAAACTG 3435

Db 1267 ATCCTTCAGCAGCTTGTCCTTGTAGACATTTACTGACGGATGGAAGAGCAGCAGCTT 1326
          | ||| |||||||||||| |||| ||| ||||| |||||| ||||||||| |||
Qy 3436 AGGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480

Db 1327 CGGGGCCCTGGAGCAGGAGGAAAAGGCCATCCCTGAGGATCATTCCTGA 1371

RESULT 7
US-09-340-620A-50
; Sequence 50, Application US/0934062OA
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREO
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-340-620A-50

Query Match   2.1%;   Score 115;   DB 4;   Length 585;
Best Local Similarity 61.5%; Pred. No. 1.5e-20;
Matches 184;   Conservative 0;   Mismatches 115;   Indels 0;   Gaps 0;

Qy 4476 CCCTCCAGCCCGCCATAGCCGTACCTTCACCTCTGGATGCCCGAGTGACATGGGTGGAGTTGTCTTTGGACAA 4535
          || ||||||||| || ||||| ||| ||||| |||||| ||||| ||||| |||| ||| ||| |||||||
Db   285 CGCCCAGCTGGGATCCAGGCCCTCCAGCGCCCGCCCTCAGCTGCCGCAGCCAAGCCAGGCCTGCACTTTAT 344

Qy 4536 GGACCAGTATCGAGACGAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTTGCTTTGGATGC 4595
          |||||||| ||||| ||| ||| || ||||| |||| ||||||| || ||| || |||||||
Db   345 AGACCAGCACCGGGCTGCGGCTTATCGGCAGGGGTCACAAACGTTGAGTGGGCTGCTGGATGC 404

Qy 4596 ACTGCATGACAGGTGTCTGAGCCAGGAGCAGTACCAGAGAGGGTGCTGGCTGTGAGAACACGAG 4655
          ||||| |||| ||| |||||||| |||||||||| ||||||| ||||||||| ||| ||| ||
Db   405 TCTGTACGGGAGAGGTCCTGACGGGATGAGCAGTACCAGCAGTGGGCGGCCGACCCCACCA 464

Qy 4656 GCCCAGCCAGATGCGGAAGCTGTTCAGCTTGAGCTTGAGGCAGTACCAGGCAGTGCGGGAAGTGCAAAGA 4715
          |||||||||||| |||||||||||||||| ||||| ||||| |||||||| ||||| ||||||||||||
Db   465 CCCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACCAGCCTGGGACTGGAACTGGAACTGTCAAGGA 524

Qy 4716 TGGACTTCTACCAAGCCCCTGAAGGAGAGACCCATCCTACCTCACCTCATTATGGAACTGTGGGAGA 4774
          |||| || ||| |||||| |||| |||| ||| || ||| |||||| || ||| ||||| ||||
Db   525 CTTGCTCCTCCAGCCCCTAAGGGAGTCCCAGTCCTACCTGGTGGGAGGAGGAACCTGGAGCGGGA 583

RESULT 8
US-09-340-620A-48
; Sequence 48, Application US/0934062OA
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREO
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281

RESULT 9
US-09-340-620A-62
; Sequence 62, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 0734-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-340-620A-48

Query Match   2.1%; Score 115; DB 4; Length 740;
Best Local Similarity 60.9%; Pred. No. 1.8e-20;
Matches 187; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy   4476 CCCTTCCAGCCCGCAATAGCCGTACCTTCACCTCTTGGATGCCCCGCAGTTGCTCACTTTGT 4535
                                             ||||||||||||||||| |||||
Db   338 CGCGCCAGCTGGGATCCAGGACCCCTCCTCAGTCGGACGCCAAGGCCAGGCCTGCACTTTAT 397

Qy   4536 GGACCAGTATCGAGAGCAGCACTGATAGCCCGGAGTGACATCGGTGGTGTTGTTCTTGGACAA 4595
           ||||||||||||||||||||
Db   398 AGACCAGCACCGGCGCTGCGCGCTTATCGCGAGGGTCACACAAACGTTGAGTGGCTGGATGC 457

Qy   4596 ACTGCATGACACAGTGCTTAGCCAGGAGCAGTACGAGAGGGTGCTGCTGAGAACACAGAG 4655
                                        ||||||||||||||||||||
Db   458 TCTGTACGGGAAGGTTCCTGACGGATGAGCAGTACCAGGCAGTGTGGGGCCCGAGCCACCAA 517

Qy   4656 GCCCAGCCAGATGCGGGAAGCTGTTTCAGCGTTGAGCCGGACCGGAAGTGCAAAGA 4715
                             |||||||||||||||||||||||||
Db   518 CCCAAGCAAAGATGTCGGAAGCTCTTCAGTTTCACACAGCCTGGAAACTGACCTGCAAGGA 577

Qy   4716 TGGACTCTACCAAGGCCCTGAAGGAGGACCCATCCTCACCTCATTATGGAACTCTCGGGAGA 4775
                   ||||||||||||||||||||||||||||||||||
Db   578 CTTGCTTCCTTCCAGGCCCTAAGGGAGAGTCCCAGTCCTTACCTGGCTGGAGAGCGGGAG 637

Qy   4776 GGGCCAGC 4782
              ||
Db   638 CTGAGGC 644

Query Match   2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 3.3e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy   4768 TGGGAGA 4774
              |||
Db   571 GAGCAGA 577

RESULT 10
US-09-340-620A-60
; Sequence 60, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 0734-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 60
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(667)
US-09-340-620A-60

Query Match   2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 4e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy   4528 CACTTTGTTGGGACCAGTATCGAGAGCAGCAGCTTGATAGCCCGAGTGACATCGGTGGGGAGGTTGTC 4587
                  ||||||||||||| ||||||||||||||||
Db   331 CACTTTGTTGGACCCAGCACAGCAGACACTCATTGCCAGGGTCACACAGAAGTGGACGGAGTG 390

Qy   4588 TTGGACAAACTGCATGGACAGGTGCTTGAGCCGGAGGAGCAGTACGAGAGGGTGCTGCTGAG 4647
           ||| |||||||||||||||||||||||||||||||||||||
Db   391 CTGGATGCGTTTGCATGGCAGTTGGCTGTCTGACTGAAGGAACAGTACCAGGCAGTTCACTGA 450

Qy   4648 AACACGAGGCCCCAGCGATGCGGAAGCTGTTCAGCGTTGAGCTTGACGCTTTCAGCTTTGT 4707
           |||||||||||||||||||||||||||||||||||||||||
Db   451 ACCACCAGCCCAAGACAAAGATGAGGAAAGCTCTTCAGCTTTTGTTTCATTCCTGGAAACCTGACC 510

Qy   4708 TGCAAAGATGGACTCTTACCAAGCCCCTGAAGGAGGACCCCCATCCTCACCTCATTATGGAAACTC 4767
                  ||||||||||| |||
Db   511 TGCAAGGACTACCCCTCTCCAGCCTTGAAGGAAAAATACATCCCTACTTGGTGATGGAACCTG 570

Qy   4768 TGGGAGA 4774
              |||
Db   571 GAGCAGA 577

Query Match   2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 4e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy   4528 CACTTTGTTGGGACCAGTATCGAGAGCAGCAGCTTGATAGCCCGAGTGACATCGGTGGGGAGGTTGTC 4587
                  |||||||||||||
Db   419 CACTTTGTTGGACCCAGCACAGCAGACACTCATTGCCAGGGTCACAGAAGTGGACGGAGTG 478

Qy   4588 TTGGACAAACTGCATGGACAGGTGCTTGAGCCGGAGGAGCAGTACGAGAGGGTGCTGGCTGAG 4647
           |||
Db   479 CTGGATGCGTTTGCATGGCAGTGTGTCTGACTGAAGGAACAGTACCAGGCAGTTCACTGA 538

Qy   4648 AACACGAGGCCCAGCGATGCGGAAGCTGTTCAGCGTTGAGCTTGACGTCCTGACC 4707
           ||||||||||||||||||||||||||||||||||
Db   539 ACCACCAGCCAAGACAAAGATGAGGAAGCTCTTCAGCTTTTGTTCCATTCCTGGAAAACCTGACC 598

Qy   4708 TGCAAAGATGGACTTCTACCAAGCCCCTGAGGGAGGACCCCATCCTCACCTCATTTATGGAAACTC 4767
              ||||||||
Db   599 TGCAAGGACTCCCTCTCCTCCAGGCCCTTGAACGGAAAATACATCCCTACTTGGTGATGGAACCTG 658

```
Qy        4768 TGGGAGA 4774
Db         659 GAGCAGA 665
```

RESULT 11
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F.G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/232,463
;   FILING DATE:
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/07/935,313
;   FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
;   FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)836-9300
;   TELEFAX: (703)683-4109
;   TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 7218 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   CLONE: pT2gpt-Fls
US-08-232-463-14

Query Match   1.1%; Score 62; DB 1; Length 7218;
Best Local Similarity 2.8%; Pred. No. 8.5e-06;
Matches 11; Conservative 230; Mismatches 145; Indels   0; Gaps   0;

Qy   23 GAAGAAACCTGGGAGCCAGCAGCCCGGGGGGCTCCACTCTCGGGTTCTGAAAGCCCATTCCCT 82
                  |||||||:||||::||| |||:: :|| ||| :||: ::||:::|||:|||||
Db   1052 GAGGGAGCTTGCGGRATYYYYYYYYYYYYYYY:::::YYYYYYYYYYYYYYYYYYYYYY 1111

Qy   83 GCTCTGCGCCTCTCCCACCCCACCTCTTCTTCAGCGCTTGCAGCCGCTCAAGGGTTGANTCTCAG 142
                  ||||:|| ||||||||:||||||||||YYYYYY:|||:YYYYYYY:YYYYYYYYYYYYYY
Db   1112 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY::YYYYYYYYYYYYY 1171

Qy   143 GAGTCCAGGAGACCCAGGAGGAGGAAGAAATCTGAGGAACACAGAACAGTGAGCGGTTGCCCAC 202
                 ||::||YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY|||:::YYYYYY:YYYYYYYYYYYYYY
Db   1172 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1231

Qy   203 ACCCCATCTCCCGTCACCACATCTCCCTCACCCTCCCTGCCTGGCCCTTGGACC 262
                  :::::||:::::::YYYYYYYYYYYYYYYYYYYYY:::::YYYYYYYYYYYY
Db   ::
```

```
Db        1232 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1291
Qy         263 CCATCCCAGGACCTCCCTATCAGCTGAACTTCTTCCAGTGTGTCTTGCAGGGCCCCTCTGGGGCT 322
                :: |:::: :||:::YYYYYYYYYYYYYYYYYYY:::: |||: ||:::YYYYYY :
Db        1292 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY:::::YYYYYYYYY 1351
Qy         323 CCTCCCTCCCTCCCGGCTGTTTCCTACCACTCCCCCTCCCCTCTATCGCGGTCGTCATCGTAGGTGCCC 382
                 :::::::|||:::::YYYYYYYY:|||:::YYYYYY ::::::YYYYYYY:::YYY ::
Db        1352 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYY:::::YYYYY:YYY:: :::
Qy         383 TGGGATTTATTAAAAACTGGGGTTCCGAA 408
                  ||:: ::::YYYYYYYY::: ||
Db        1412 YYYYYYYYYYYYYYYYYYYYYYYYGTA 1437
```

RESULT 12
US-09-016-434-208
; Sequence 208, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 208:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 257 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: SINTBST01
;   CLONE: 1499408
US-09-016-434-208

Query Match   1.1%; Score 59.4; DB 4; Length 257;
Best Local Similarity 57.5%; Pred. No. 5.2e-06;
Matches 122; Conservative   0; Mismatches 89; Indels   1; Gaps   1;

```
Qy        1987 TATTTCTTACAGATATTTTCACAGATGAAAAGGCAAGCAAATTAGAGCGCTTTAGGTTGGTTCAAA 2046
                |||||||||| ||||||| ||||||||||||||||||||||||||||||||||
Db           2 TATTTCTTACAAGTATTTTCCGGGATGAGAGAGGGCCAGAGGGCCGAGCGCCTACCCGCTTCGTGAAG 61
Qy        2047 TCAAACAAAAGAGCTCTGGGCCCTGTGTTCTTGTGCCCTGGGTGTCTTGCCCTGGCCCTGCACT 2106
                |||||||||||| ||||||||||||||||||||||||||||||||||||||||
```

Db   62 GAGAACGACGACGCTGTTCGCGCGTCGTGTCTTCGTGCCCTTCGTTCGTGTCGTGTGCACC 121
          |||| ||| | | |||| | ||| | | || || | | ||| ||||| |||| ||
Qy 2107 TGCCTGATGCAGCAGCAGATGAAGCGGAGGAAGAAAAACTCACACTGACTTCCAAGACCACCACA 2166

Db   122 GTGCTG--CGCCANANTGNGAGCTCGGTCGGGAGACCTGTCGGGCACCTCCAAGACCACCACG 180
          || ||| | || | | |||| | | | | || | | | |||||||||||||||
Qy 2167 ACCCTCTGTGTCTACATTACCTTGCCCAGGCTCT 2198

Db   181 TCAGTGTACCTGCTTTTTCATCACCAGCGTTCT 212

RESULT 13
US-09-904-615-43

Sequence 43, Application US/09904615
Patent No. 656325

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 43
; LENGTH: 3116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-43

Query Match   50.9%; Score 57.2; DB 4; Length 3116;
Best Local Similarity   1.1%;
Matches 162; Conservative   0; Mismatches 153; Indels   3; Gaps   1;

Qy 3023 GCCCTCGCGTGCCTCCTCGGAGACACCCTGCGTTGGCGTGGTGGCCCTCACACGTCGAGGACT 3082
          |||| || | || | ||| || | ||| || ||||| || || ||| || |||| ||||||
Db 1036 GCCACTGCCACCACTTGAGGAGCTGGACTTGTCTTAACANATCAAATTGATGAGGAGGGCA 1095

Qy 3083 GCAAGGACCTTGCCCTTTGGGCTGAGAGCCAACCAGCCTGACCGAGCTGGACCTGAGCT 3142
          || || || | | ||| | |||| || || || ||||| ||| ||||||| ||||
Db 1096 CCAAGGCCGCTGATGAGGGGCCCTTGAGGGGGAAATGGATAGCTAAAGGGCTGGACCTCAGTC 1155

Qy 3143 TCAATGTGCTCACGGATGCTGGAGCCAAAACACCTTTGCCAGAGACTGAGACACCCGACCT 3202
          ||| ||| | | || ||| ||| || | ||| || | | | ||| ||| | || |||
Db 1156 ACCTTCTCTGTGAACAGCTCCACCTTGGCCCTTGCTTACTCACAGACTAAGCCAGATGACCT 1215

Qy 3203 GCAAGCTACAGCGACACTGCGCAGCTGGTCAGCTGTGTGGCCTCACGTCTGACTCGTCGCCAGGACC 3262
          || ||||| || | | |||| | || || | ||| ||||| || | | |||| ||| ||||||
Db 1216 GC---CTGCAGAGCCTCAGACTGAACAGCAGTATCTCGGTGAATGTCGGTTCGGTTGCTGCCACC 1272

Qy 3263 TGGCCTTCGTTGCTTAGTTGCTGCCAGCCCCGAAGGAGGGCTAGACCTGCAGCAGAGAACAACC 3322
          |||| | | | || | | | ||| |||| | |||| | | | | | |||||| | | |
Db 1273 TTTCTGAGGCCTTCTCCAGGGCTGCCACCAGCGTGACTTGAGCCACACAGA 1332

Qy 3323 TGGATGAGCTTGCGCTGC 3340
          | |||| || || || ||
Db 1333 TTGGAGACGCTCGGTGTCC 1350

RESULT 14
US-08-594-031-79

Sequence 79, Application US/08594031
Patent No. 5783182

; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-594-031-79

Query Match   62.8%; Score 49; DB 1; Length 337;
Best Local Similarity   0.9%;
Matches 76; Conservative   0; Mismatches 45; Indels   0; Gaps   0;

Qy 5317 ATTTGTCTAGCAGGCGCTTTCCAGTTTTTACCAGAAAACCCCTATAAATTAAAATTTTTTA 5376
          |||| | || ||| || |||| | ||||| |||| | |||| ||||||||||
Db   215 ATTTATTGATAATGTTTTGCTACTCCCTGTCAGRCAATGGCTATAAACTGAAATTAGGCAGT 274

Qy 5377 CTTAAAATTTAAGAATTAAAAAAAAAATTACAAAAAAAAAGAAAAAATTGAAAAAAAAGGAATAAGAGAG 5436
          ||| ||||||| | |||||||||||| |||||||||||||||| || |||| ||| ||| || |
Db   275 CTTAAAAAAAAAAGAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAGAAAAGAAAAAAAC 334

Qy 5437 T 5437
          |
Db   335 T 335

RESULT 15
US-09-465-558-39

Sequence 39, Application US/09465558
Patent No. 6436657

; GENERAL INFORMATION:
; APPLICANT: Morakinyo, Layo O.
; APPLICANT: Orozco Jr, Emil M.
; TITLE OF INVENTION: TETRAHYDROFOLATE METABOLIC ENZYMES
; FILE REFERENCE: BB1322 US NA
; CURRENT APPLICATION NUMBER: US/09/465,558
; CURRENT FILING DATE: 1999-12-17
; EARLIER APPLICATION NUMBER: 60/112,734
; EARLIER FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 39
; LENGTH: 1878

```
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-09-465-558-39

Query Match        0.9%;    Score 48.6;   DB 4;   Length 1878;
Best Local Similarity  52.8%;   Pred. No. 0.013;
Matches  105;   Conservative   0;   Mismatches  94;   Indels   0;   Gaps   0;

Qy    5237 AGATCTACAAGAGCAAGCCAACAGAGTAAAGTGGBAGGBAGTTTATTCAGAAAATAAAGG 5296
           ||||| ||   || |  |   | | ||||   |   |  || |    ||| |  | ||
Db    1672 AGATCCAATGAATGAAGCGCCACCATTATGTAAAAGGAATCCAAGCAATTTTACAGAATATGGG 1731

Qy    5297 AGTATCACAGCCTCTTTTAGAATTTGTCTAGCAGGGCTTTCCAGTTTTTACCAGAAAACCCC 5356
           || |||| |  | ||| ||  || |  | |  |  || || ||     | |||| |||
Db    1732 AACTTGTCGATAGTTTCTTATTGCCAGGGCACTACTGTGTAAGAATGGCTTCGGCTGATATGCTA 1791

Qy    5357 TATAAAATTAAAAAATTTTTTACTTAAATTTAAGAATTAAAAAATACAAAAAGAAAAAAT 5416
           |||| |||  || || |||||| |||  ||  || ||| |  || || | |||||||
Db    1792 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1851

Qy    5417 GAAAATAAAGGAATAAGAA 5435
           ||||  ||  ||| || ||
Db    1852 AAAAAAAAAAAAAAAAAAA 1870
```

Search completed: January 29, 2004, 13:11:27
Job time : 293.391 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 13:44:42 ; Search time 35.197 seconds
                                            (without alignments)
                                            3904.448 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWGRLACYLEFLKKEE.........HLIMELWEKGSKKGLLPLSS 1429

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_76:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	DB ID	Description
1	5969	79.2	1192	2	T17255	hypothetical prote
2	882	11.7	1111	2	A59000	mater protein [imp
3	465.5	6.2	483	2	S27880	Nasopressin recept
4	361.5	4.8	456	2	S20597	ribonuclease inhib
5	358	4.8	461	2	A31858	ribonuclease-angio
6	354.5	4.7	456	2	A31858	ribonuclease inhib
7	345	4.6	1130	2	A48843	MHC class II trans
8	206	2.7	1075	2	T31668	hypothetical prote
9	202	2.7	1004	2	T31665	hypothetical prote
10	198	2.6	1232	2	A55478	hypothetical prote
11	196.5	2.6	1447	2	T42628	neuronal apoptosis
12	180	2.4	533	2	T52063	ran GTPase-activat
13	166.5	2.2	1120	2	JC7765	mitotic spindle as
14	159.5	2.1	506	2	A45841	T-complex-associat
15	155.5	2.1	1199	2	T47442	disease resistance
16	153.5	2.0	1027	2	F96662	hypothetical prote
17	152	2.0	1027	2	T46296	hypothetical prote
18	152	2.0	3511	2	A59295	unconventional myo
19	151	2.0	1220	2	T52347	disease resistance
20	150.5	2.0	1220	2	T48928	disease resistance
21	149	2.0	631	2	C89243	protein F28C1.3 [i
22	149	2.0	631	2	T21471	hypothetical prote
23	148	2.0	1217	2	T52348	disease resistance
24	147.5	2.0	793	2	S54772	mammary gland fact
25	147.5	2.0	1214	2	T47438	disease resistance
26	147	2.0	1017	2	H96663	hypothetical prote
27	147	2.0	1784	2	T10532	hypothetical prote
28	143.5	1.9	4302	2	A38971	gag-pol polyprotei
29	142	1.9	1215	2	T00364	polycystic kidney hypothetical prote
30	142	1.9	1839	1	OYBYK	adenylate cyclase
31	139.5	1.9	545	2	T52068	RAN GTPase-activat
32	139.5	1.9	1189	2	T52346	disease resistance
33	138.5	1.8	1031	2	G96663	hypothetical prote
34	138	1.8	1010	2	T36383	probable large ATP
35	137	1.8	1226	2	T45788	probable disease r
36	136.5	1.8	794	2	G02317	transcription acti
37	136.5	1.8	3335	2	H81702	adherence factor T
38	135.5	1.8	1422	2	B71437	probable resistanc
39	135	1.8	915	2	T09575	smoothelin - human
40	135	1.8	1724	2	T18343	P-glycoprotein - S
41	134	1.8	889	2	H96606	hypothetical prote
42	133	1.8	1184	2	H71436	hypothetical prote
43	133	1.8	1301	2	D85188	disease resistance
44	133	1.8	1453	2	G96613	hypothetical prote
45	132.5	1.8	805	2	AH2731	ATP-dependent prot

ALIGNMENTS

RESULT 1
T17255
hypothetical protein DKFZp58601822.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17255
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17255
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1192 <KOE>
A;Cross-references: EMBL:AL117470
A;Experimental source: adult uterus; clone DKFZp58601822
C;Genetics:
A;Note: DKFZp58601822.1

Query Match   79.2%; Score 5969; DB 2; Length 1192;
Best Local Similarity   96.2%; Pred. No. 0;
Matches 1147; Conservative   1; Mismatches 0; Indels 0; Gaps 1;

Qy   282 LLLLQRPHPRSQDPLVKRSWPDYVEENRGHLLIEIRDLFGPGLDTQEPRIVILQGAAGIGKS 341
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 LLLLQRPHPRSQDPLVKRSWPDYVEENRGHLLIEIRDLFGPGLDTQEPRIVILQGAAGIGKS 60

Qy   342 TLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSR 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSR 120

Qy   402 PERLLFILLDGGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTT 461
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 PERLLFILLDGGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTT 180

Qy   462 ALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPW 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPW 240

Qy   522 VSWLACTCLMQQMKRXEKZLTLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQ 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 VSWLACTCLMQQMKRXEKZLTLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQ 300

Qy   582 KKTLFSPDDLRKHGILDGAIISTFLKQMGILQEHPIPLSYSFIHLCFQEFPAAMSYVLEDEK 641
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KKTLFSPDDLRKHGILDGAIISTFLKQMGILQEHPIPLSYSFIHLCFQEFPAAMSYVLEDEK 360

Qy   642 GRGKHSNCIIDLEKTLEAYGIHGLFGASTTRFLLGILLSDEGGEREMENIFHCRLSQGRNLM 701
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GRGKHSNCIIDLEKTLEAYGIHGLFGASTTRFLLGILLSDEGGEREMENIFHCRLSQGRNLM 420

Qy   702 QWVPSLQLLLQPHSLESLHCLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFS 761
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2
A59000
mater protein [imported] - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: A59000
R;Tong, Z.B.; Nelson, L.M.
R;Reference number: A59000; MUID:99360614; PMID:10433232
A;Title: A mouse Gene encoding an oocyte antigen associated with autoimmune premature ov
A;Title: Endocrinology 140, 3720-3726, 1999
A;Accession: A59000
A;Molecule type: DNA
A;Status: preliminary
A;Residues: 1-1111 <KUR>
A;Cross-references: GB:AF074018; NID:g5802697; PIDN:AAD51762.1; PID:g5802698
C;Genetics:
A;Gene: Mater

Query Match 11.7%; Score 882; DB 2; Length 1111;
Best Local Similarity 26.1%; Pred. No. 2.4e-49;
Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

```
Db   421  QWVPSLQLLLQPHSLESLHCLYETRNKTFLTQVVAHFBEMGMCVETDMELLLVCTFCIKFS  480
Qy   762  RHVKKLQLLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKBLDLSGNSLS  821
Db   481  RHVKKLQLIEGRQHRSTWSPTMVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLS    540
Qy   822  HSAVKSLCXTLRRPRCLLETLRLAGCGLLTAEDGKDLAFGLRANQTLTELDLSFNVLTDAG  881
Db   541  HSAVKSLCXTLRRPRCLLETLRLAGCGLLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAG  600
Qy   882  AKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSRASPSLKELDLQQNNLDDVGVRL  941
Db   601  AKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSRASPSLKELDLQQNNLDDVGVRL  660
Qy   942  LCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTG  1001
Db   661  LCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTG   720
Qy  1002  EMSNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSSPEVVPVELLCVPS  1061
Db   721  EMSNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSSPEVVPVELLCVPS  780
Qy  1062  PASQGDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMRE  1121
Db   781  PASQGDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMRE   840
Qy  1122  AVTVEIEFCVWDQFLGEINPQHSMWVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSL  1181
Db   841  AVTVEIEFCVWDQFLGEINPQHSMWVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDTSL  900
Qy  1182  FQVAHFKEEGMLLEKPARVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRV   1241
Db   901  FQVAHFKEEGMLLEKPARVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRV   960
Qy  1242  HPEEVTFHLYLIPSDCSIR                                           1260
Db   961  HPEEVTFHLYLIPSDCSIRKAIDDLEMKFQFVRIHKPPPLTPLYMGCRYTVSGSGSGMLE  1020
Qy  1261  ---KELELCYRSPGEDQLFSEFYVGHLGSSGIRLQVYDKDKDETLVWEALVKPGDLMPATTL 1317
Db  1021  ILPKELELCYRSPGEDQLFSEFYVGHLGSSGIRLQVKDKDETLVWEALVKPGDLMPATTL  1080
Qy  1318  IPPARIAVPSPLDAPQLLHFVDQYREQLIARVISVEVVLDKLHCQVLSQEEQYERVLAENT 1377
Db  1081  IPPARIAVPSPLDAPQLLHFVDQYREQLIARVISVEVVLDKLHGQVLSQEEQYERVLAENT 1140
Qy  1378  RPSQMRKLFSLSQSWDRKCKDGLIYQALKETHPHLIMELWEKGSKKGLLPLSS         1429
Db  1141  RPSQMRKLFSLSQSWDRKCKDGLIYQALKETHPHLIMELWEKGSKKGLLPLSS         1192
```

```
Qy   162  SSPDHESPSQOESPNAP---------TSTAVIGSWGSPPQPS----IAPR----EQEAPGTQWPLD  209
Db    28  TSPENDSKSIQKDQGPEEQTSESTMG----PPEKDSKAILKARGLEEQKSESTMSPSE      83
Qy   210  ETSGIYYTEIREREREKSEKGRPWAAAVGTPPQAHTSLQPHHIPWEPPSVRBSLCSTWPW  269
Db    84  NVSRAILKDSGSEEVEQASERK------MTSPENDSKSIQKDQGPEEQTSETLQS----  133
Qy   270  KNEDPNQKFTQLLLQRPHPRSQDPLVRSWPDVVEERGHLIE--------------------  313
Db   134  KEED---EVTE-----------ADKQNGGD-LQDYXAHVIAKFDTSVDLHYDSPEMK  175
Qy   314  -IRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDRFQHVFVFSCREL  372
Db   176  LLSDAFKPYQKTFQPHTIILHGRPGVGKSALARSIVLGWAAQGKLF-QKMSFVIFPSVREI  234
Qy   373  AQSKVVSLAELIGKDGTATFPAPIRQILSRPERLLFILDGVDBPGMWLQEPSSELCLHWSQ  432
Db   235  KWTEKSSLAQLIAXECPDSWDLVTKIMSQPERLLFVIDGLDDMDSVLQHDDMTLSRDWKD   294
Qy   433  PQPADALLGSLLGKTILPEASFLITARTIALQNLIPSLEQARWVEVLGFSESSSRKEYFYR  492
Db   295  EQPIYILMYSLLRKALLPQSFLIITTRNTGLEKLKSMVVSPLYILVEGLSASRRSQLVLE   354
Qy   493  YFTDERQAIRAFRLVKSNKBIWALCLVPWVSWLACTCLMQQMKRKEKLTLITSKTTT---  548
Db   355  NISNESDRIQVFHSLIENHQLFDQCQAPSVCSLVCEALQLQKGKRCTLPCQTLTGLYA    414
Qy   549  TLCLHYL------AQALAQQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIIST  603
Db   415  TLVFHQLTLKRPSQSALSQEEQITLVGLCMMAARGVWTMRSVFYDDDLKNYSLKESEILA   474
Qy   604  FLKNGILLQS--HFIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKH------SNCIIDLEK  655
Db   475  LFHMNILLQVGHNSBQCTVFSHLSLQDFFAALYVVLEGGLEEWNQHFCFIENQRSIMBVKR  534
Qy   656  TLEAYGIHGLFGASTTRFLLGLLSDBEGEREMENIFHCRL---SQGRNIMQWVPSLQLLQP  713
Db   535  -TDDT----RLLGMK--RFLFGLMNKDILKTLEVLFEYVIPTVEQKLQHHV---SLIAQQ  585
Qy   714  ------HSLESLHCLYETRNKTFLTQVVAHFBEMGMCVETDMELLLVCTFCIKFSRHVKK  766
Db   586  VNGTSPMDIDAFYCLFBSQDEEFVGGALKRFQBWVLLINQKMDLKVSSYCLKHCQNLKA   645
Qy   767  LQL------IEGRQHRSTWSPTMVVL-FRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLS  821
Db   646  IRVDIRDLLSVDNTLELCPVVTVQETQCKPLLMEWWGNFCSVLGSLRNLKELDLGDSILS   705
Qy   822  HSAVKSL                                                       828
Db   706  QRAMKIICLELRNQSCRIQKLTFKSAEVVSGLKHLWKLLFSNQNLKYLNLGNTPMKDDDM  765
Db   766  KLACEALKHPKCSVETLRLDSCELTIIGYEMISILLISTTRLKCLSLAKNRVGVKSMISL  825
Qy   829  ---CKTLRRPRCLLETLR                                            843
Db   826  GNALSSSMCLLQKLLDNCGLTPRASCHLLVSALFSNQNLTHLCLSNNSLGTBGVQOLCQF  885
Qy   844  LAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQR                 888
Db   886  LRNPECALQRLILNHCNIVDDAYGFLAMRLAANNTKLTHLSLTMNPVGDGAMKLLCEALKE  945
Qy   889  LRQPSCKLQR                                                    898
Db   946  PTCYLQELELVDCQLTQNCCEDLLACMITTTKHLKSLDLGNNALGDKGVITLLCEGLKQSS 1005
Qy   952  KLIRLGLDQTTLSDEMRQELRALEQEKPQL                                981
Db  1006  SLRRLGLGACKLTSNCCEALSLAISCNPHL                               1035
```

RESULT 3
S27880

C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
R;Herrera, V.L.; Ruiz-Opazo, N.
submitted to the EMBL Data Library, February 1992
A;Reference number: S27880
A;Accession: S27880
A;Molecule type: mRNA
A;Residues: 1-483 <HER>
A;Cross-references: EMBL:M85183; NID:g202805; PIDN:AAA03623.1; PID:g202806
C;Genetics:
A;Gene: AVP

Vasopressin receptor - rat
A;Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.

Query Match   6.2%; Score 465.5; DB 2; Length 483;
Best Local Similarity 30.6%; Pred. No. 1.3e-22;
Matches 161; Conservative 74; Mismatches 186; Indels 105; Gaps 18;

Qy 534 MKRKEKLTLTSKTTTTLCLHYLAQALQ-AQPLGP----QLRDLCSLAAEGIWQKKTLFSP 588
               ||||| :::: ||| : : ||||| |: |||| |:|||| :||
Db   1 MELGRDLSRTSKTTTSVLLFITSMLKSAGTNGPRVQGELRMLCRLAREGILKHIQAQFSE 60

Qy 589 DDLR----KHGLDGAIIST-FLKM----GILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEK 641
          |||| : :||| :: :|:   ::| :|| : ||||| | : : ||: :|
Db   61 KDLERLKLKLQGSQVQTMFLSKKELPGVLE---TVVTYQPIDQSFQEFLAALSYILDAE- 116

Qy 642 GRGKHSNCIIDLEKTLEAYGIHGLPG--ASTTRFLLGLLSDEGEREMENIFHC----RLS 695
          |||| | :| | ||: | ||| : ||::|| |||
Db 117 --GAPGNSAGSVQMLLNSDA--GLRGHLALTTRFLLFGLLSTERIRDIGNHFGCVVPGRVK 172

Qy 696 QGRNLMQWVPSLQLLLQP---------------------------HSLESLHCLY 723
          || |: : :|||||| ||::|||:|
Db 173 Q--DTLRWV---QGSQPKVATVGAEKKDELKDEAAEEEEEEEEEELNFGLELLYCLY 227

Qy 724 ETRNKTFLTQVMAHFEEMGMCVE---------TDMELLVCTFCIKFSRHVKKLQLI------E 771
          | |: : | : | | : |::| | : |: |
Db 228 ETQEDDFVRQALSSLPE--MVLRVRLTRMDLEVLSYCVQCCPDGQALRLVSCGLVAAKE 285

Qy 772 GRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKT 831
          |: :| | || | :: : ::|||: | | | ||||:|||||| :||| ||||
Db 286 KKKKKKSP--------------------MNRLKGSQSTGKQPPASLLRPLCRA 318

Qy 832 LRRPRCLLETLRLAGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQ 891
          || : ||| ||| ||| :||| || |:||| ||:|||||||: | |||| ||||||
Db 319 MITQQCGLSILTLSHCKLPDAVCRDLSRALKVAPSLRELGLLQNRLTERAGLRLLSQGLAW 378

Qy 892 PSCKLQRLQVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPAC 951
          ||||||| | | || :|||| | |||||||| |||||||||||| ||| |
Db 379 PKCKVQTLRIQMPGI--QEVIHYLVIVLQQSPVLTTLDLSGCQLPGTVVEPLCSALKHPKC 437

Qy 952 KLIRLGLDQTTLSDEMRQELRALEEQEKPQLLIFSRRKPSVMTPTEG 997
          |:|||||||||||||||||| | |||||:|:||| :| |
Db 438 GLKTLSLTSVELTENPLRELQAVKTLKPDLAIHSKLGTHPQPLKG 483

RESULT 4
S20597

Ribonuclease inhibitor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
A;Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distributi
Biochim. Biophys. Acta 1129, 335-338, 1992
A;Reference number: S20597
A;Accession: S20597
A;Molecule type: mRNA
A;Residues: 1-456 <KAW>
A;Cross-references: EMBL:X62528; NID:g57670; PIDN:CAA44388.1; PID:g57671
A;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homolog

F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match   4.8%; Score 361.5; DB 2; Length 456;
Best Local Similarity 35.5%; Pred. No. 7.9e-16;
Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;

Qy 792 VTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGGLTA 851
          | |: ||||: | | | ||||:||||| :|||| |||| || : ||| ||| |||
Db 793 ITSANCKDLCDVVASKASLQELDLGSNKLGNTGIAALCSGLLPSCRLRTLWMDCDVTA 851

Qy 852 EDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQVSCGLTS--- 908
          ||||| ||:|| | ||:|||| | ||| | || :: |: : | |
Db 852 EGCKDLCRVLRAKQSLKELSLAGNELKDEGAQLLCESLLEPGCQLESLWVKTCSLTPAASC 325

Qy 909 ---DC-------CQDL 914
                                                         || ||||
Db 326 PHFCSVLTKNSSLFELQMSSNPLGDSGVVELCKALGYPDTVLRVLWLGCDVTDSGGCSSL 385

Qy 915 ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRAL 974
          ||||:|||||||||| |||||||||||||||||: | |||:||||||||||||||||| |
Db 386 ATVLLANRSLRELDLSNNCMGDNGVLQLLSSLKQPSCILQOLVLIVDIYWTDEVEDQIRAL 445

Qy 975 EQEKPQLLIFS 985
          || :| |||
Db 446 EEERPSLRIIS 456

RESULT 5
A31858

Ribonuclease-angiogenin inhibitor - human
N;Alternate names: ribonuclease inhibitor, placental
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000
R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
Biochemistry 27, 8545-8553, 1988
A;Title: Primary structure of human placental ribonuclease inhibitor.
A;Reference number: A31858; MUID:89118269; PMID:3219362
A;Accession: A31858
A;Molecule type: mRNA
A;Residues: 1-461 <LEE>
A;Cross-references: GB:M22214; NID:g162260; PIDN:AAA59130.1; PID:g307040
R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
EMBO J. 7, 4151-4156, 1988
A;Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) disclose
A;Reference number: S02012; MUID:89210799; PMID:3243277
A;Accession: S02012
A;Molecule type: mRNA
A;Residues: 1-422, 'SE', 425-461 <SCH>
A;Cross-references: EMBL:X13973; NID:g35843; PIDN:CAA32151.1; PID:g35844
A;Note: part of this sequence, including the carboxyl end of the mature protein, was con
R;Crevel-Thieffry, I.; Cotterill, S.; Schuller, E.
Biochim. Biophys. Acta 1122, 107-112, 1992
A;Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibit
A;Reference number: S23933; MUID:92338217; PMID:1633192
A;Accession: S23933
A;Molecule type: protein
A;Residues: 174-195 <CRE>
R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchide, K.; Kishi, K.
Arch. Biochem. Biophys. 312, 421-428, 1994
A;Title: Purification and characterization of human brain ribonuclease inhibitor.
A;Reference number: S48636; MUID:94311593; PMID:8037455
A;Accession: S48636
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-14 <NAD>
R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24374
A;Accession: T47188
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <AAA>

A;Cross-references: EMBL:AL161967
A;Experimental source: adult testis; clone DKFZp434K249
C;Genetics:
A;Gene: GDB:RNH
A;Cross-references: GDB:125274; OMIM:173320
A;Map position: 11p15.5-11p15.5
A;Note: DKFZp434K249.1
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: blocked amino end; duplication
F;57-81/Domain: leucine-rich 57-residue repeats
F;34-433/Region: leucine-rich alpha-2-glycoprotein repeat homology
F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;174-195/Region: inhibitory
F;228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Qy   Query Match   4.8%;   Score 358;   DB 2;   Length 461;
      Best Local Similarity   37.9%;   Pred. No. 1.4e-15;
      Matches 86;   Conservative 30;   Mismatches 91;   Indels 20;   Gaps 3;

Qy   768 QLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQI-----------LFSVLKVTRNLKELDLSGN 818
          ||||||||||||||||||||||| ||| :||| |:|||||||||| |||||
Db   246 ELCPGLLHFPSSLRTL-----WI-----WECGITAKGCGDLCRVLRAKESLKELSLAGN 294

Qy   819 SLSHSAVKSLCKTLRRPCLLETLRLAGGLTAEDCKDLAFGLRANQTILTELDLSFNVLT 878
          |||:|| |||:|||| ||:||| | : | |: |:| | ||| | :||:||||| |||
Db   295 ELGDEGARLLCETLLEPGCQLESLWVKSCSFTAACCSHFSSVLAQNRFLLELQISNNRLE 354

Qy   879 DAGAKHLCQRLRQPSCKLLVSGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVG 938
          |:| |:|| || | | :| |: | |:||| |||:|:| ::|||||| |:| ||
Db   355 DAGVRELCCGLGQPCGSVLRVLWLADCDVSDSSCSSLAATLLANHSLRELDLSNNCLGDAG 414

Qy   939 VRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALREQEEKPQLLIFS 985
          ||:|||||:|||:| |:|:|| ||:::|:|| :|::|::|:||||||
Db   415 ILCQLVESVRQPGCLLEQLVLYDIYWSEEMEDRLQALEKDKPSLRVIS 461

RESULT 6
A31857
ribonuclease inhibitor, hepatic - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
A;Accession: A31857; A35830
R;Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
Biochemistry 27, 8537-8544, 1988
A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals th
A;Reference number: A31857; MUID:89118268; PMID:3219361
A;Accession: A31857
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-456 <HOF>
R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
Biochemistry 29, 8827-8834, 1990
A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea
A;Reference number: A35830; MUID:91104783; PMID:2271559
A;Accession: A35830
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 82-456 <VIC>
A;Cross-references: GB:M58700; NID:g164638; PIDN:AAA63448.1; PID:g164639; GB:J02925
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: liver
F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Qy   Query Match   4.7%;   Score 354.5;   DB 2;   Length 456;
      Best Local Similarity   34.3%;   Pred. No. 2.3e-15;
      Matches 108;   Conservative 35;   Mismatches 101;   Indels 71;   Gaps 8;

Qy   752 LVCTFCIKFSRHVKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLK 811
          |::| ||| |:|||| |||||||||||||||||||||||||||||||||||||:
Db   127 LLCEGLLDPQCHLEKLQL---------------------EYCRLTAASCEPLASVLRATRALK 168

Qy   812 ELDLSGNSLSHSAVKSLCKTLRRPCLLETLRLAGGLTAEDCKDLAFGLRANQ-TLTEL 870

```
Qy   438  -ALLGSLLGKTILPEASFLITARTTALQNIIPSLEQA-RWVEVLGFSESSRKEYFRYFT  495
          .::. :.. :..:: ||.:.|:: : .:..|.. .  .:|.|.|||||||||||||.|:
Db   526  RGLLAGLFQKLLRGCTLLLTARPRG--RLVQSLSKADALFELSGFSMEQAQAYVWRYFE  583

Qy   496  D----ERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTTLC  551
                 |:.. |.|..: .:.. :.   . . .|..:.|:. :.||.|:.|..|:.:.
Db   584  SSGMTEHQD-RALTLLRDRPLLISHSHSPTLCRAVCQLSRALLELGEDAKLPS-TLTGLY  641

Qy   552  LHYLAQALQAQPLGPQLRDICSLAAE-GIWQKKTL---FSPDDLRKHGLDGAIISTFLK  606
          :. |.:.|::.||| .|.:. .|.:: ||:.:: .   |||||||||||||||||| ||
Db   642  VGLLGRAALDSPPG-ALAELAKLIAWELGRRHQSTLQEDQPPSADVR------      691

Qy   607  MGILQEHP----IPLSY-SFIHLCF-QEFFAAMSVVLEDE------KGRGKHSNCI  650
          :||||||:    |||.| |||:|||.|||||||||| |.:       ||| |:||:
Db   692  KGLVQHPPRAAESELAPPSFLLQCFLGALWLALSGEIKDKELPQYLALTPRKKRPYDNWL  751

Qy   651  IDLEKTLEAYGIHGLFPGASTRFPLLGLLSDEG------EREM                 686
          |::.|. .||.:.||:|| ||||.||||| :.        :..
Db   752  BGVPRPLA-----GLIFQPPARCLGALIGPSAAAASVDRKQKVLARYLKRLQPGTLRARQL 806

Qy   687  ENIFHC-----RLSQGRNLMQWVP------SLQLILQPHSLE---               717
          ||:|.|     ||||||||||||.      |||::|||||:|
Db   807  LELLHCAHEABEAGIWQHVVQELPGRLSFLGTRLTPPDAHVLGKALEAAGQDFSLDLRST  866

Qy   718  ------SLHCLYETRNKTFLTQVMAHFEEMGMCVETDM-----ELLVCIFPCIKF     760
                |.||||||||||||||||||||||||||||||    |||.|.|||:||
Db   867  GICPSGLGSLVGLSCV--TRFRAALSDTVALWESLRQHGETKLLQAAEEKFTIEPFKAKS  924

Qy   761  SRHVKKL-QLIEGRQHRSTWSPTMVVL-----FRWVPVTDAY-WQILFSVLKVTR     808
          ||||||| |||||||||||||||||||     |||||||||| |||||||||||:
Db   925  LKDVEDILGKLVQTQRTRSSSEDTAGELPAVRDLKKELFALGPVSGPQAFFKLVRILTAFS 984

Qy   809  NLKELD---LSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGGLITAEDCKDLAFGLRANQ  865
          ||||:|   |||||||||||||||||||||||||||||||.||||||||||||||||||.
Db   985  SLQHLDLDALSENKIGDBGVSQLSATF--POL------K                     1015

Qy   866  TLTBLDLSFNVLTDDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLK 925
          :||:|||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db  1016  SLETLNLSQNNITDLGAYKLAEALPSLAASLLRLSLYNNCICDVGAESLARVLPDMVSLR  1075

Qy   926  ELDLQQNNLDDVGVRLLCEGLR                                       947
          :|||.|||:||||||||||||:
Db  1076  VMDVQYNKFFTAAGAQOLAASLR                                     1097
```

RESULT 8
T31668
hypothetical protein COS1.5 - sea squirt (Ciona intestinalis)
C;Species: Ciona intestinalis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C;Accession: T31668
R;Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z21050
A;Accession: T31668
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1075 <BIR>
A;Cross-references: EMBL;Z80904; PIDN;CAB02589.1
C;Genetics:
A;Introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1

Query Match   2.7%;
Best Local Similarity   18.5%;
Matches   209;   Conservative 152;   Mismatches 347;   Indels 422;   Gaps 49;
                  Score 206;   DB 2;   Length 1075;
                  Pred. No. 4.8e-05;

```
Qy   301  PDYVEENRGHLIEIRDLFGPGLDT------QEPRIVILQGAAGIGKST             342
          |.: :.|.. ::||.:|.|.|.||       ||.|:||.|.|||||||
Db    37  PEPAPVNVQPLVTIDQLFDKALENAQRYTEDESAEYAKWIERHANTVWVVGPPSGVGKT   96

Qy   343  LARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTA------TPAP   394
          | :.||:.. ::..:.|.|...||.. :::.. :|:.|.||:..:.        |.|
Db    97  LLKMMVKQILKHEFLPDT-EYIFFINVKDIDFNKEMTLLEFLTTNSRVKVNYTEEESKAL 155
```

```
Qy   395  IRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHHWSQPQPADALLGSSLLGKTILPEASF 454
          |:.||::|..:|.|.||.||.|.||.||||||||||||||||||||||||||||||||||||
Db   156  ITFLHNNPNVAIF-PDGLDEASTNEFARIPHI-CKLDGKSKPVD-IMKNLFNLTLLPKAKI  213

Qy   455  LITARTTALQNIIPSLEQARWVEVLGFSESSRKEYFRYFTDERQAIRAFRLVKSNKELW   514
          ||:.|.|..:.:.|..|..:||.|.|.....|::.|.:.:..|| ||.|.|:|::|..:
Db   214  VVTSTLHQMYKLHPDYRPTSIFEVLGLLEEAXNNLGTQLCGEKYPAIK--KILDQQPNIA  271

Qy   515  ALCLVPWVSWLACTCLMQQMRKEKLTLTS----KTTTTLCLHYLAQALQAQPLGPQLRD   570
          .|||.|:.:.||||||||||||||||||:    |:|::|||||||||||||||||||||
Db   272  HLCYLPINFLILVFCLLSN-EGSSDLKTMTQVLIFSMRFVELSHLKGSEVPLDKVGAEMVK 330

Qy   571  LCSLAABGIWQKKTLFSPDDLRKHGLDGAIISTF------LKMGILQEHPIPLSYS    620
          |.||.:.|::|..:.::|||||||||||||||:.       :||||||||:|||.|
Db   331  LARLAYKGLIQQRKLVFEKTDFDDVKLADEMVTNFFHTYVDKSSGIRMKILEGN--KRSY- 387

Qy   621  FIHLCFQEFFAAMSYVL--------------------------EREM            637
          |:||..|||||||.|:|                          ::..
Db   388  FTHLIWQEFYAAVYLMLFVSYRBFEQLKPIFKDAQWKRVVGPMFGICNPPAYKQLKLVFP  447

Qy   638  -----EDEK------------GRG------                              644
               ::.|            ||.
Db   448  ATMIKDYBEKKELMVPMMMESLWSARGEDLIRRFGWLHEYNDDESSKKFEDVLPVGLKMD  507

Qy   645  --KHSNCIIDLEKTLEAY-----------------GIHGLFGASTRFLLG-         676
            ||:||.|||:|.||:|                 ||:||||||||||||
Db   508  APKHLSEVKDLVYALKSFTKPHKLRLDSYETTTEVLETLLRGVHGT-TTITRFVINNI   566

Qy   677  ------LLSDEGEREMENI---------FHCRLSQGRNLMQWVPSLQLILQPHSLESL   719
                ||||:.:|:.||:         |||||||||||||||||||||||||||||||
Db   567  EMKDSILMELLILHLDAMEELRFIRVTNLSPYMERLSNAIN--QRSNKIQLVLLKGH---KL 621

Qy   720  HCLYETRNKTFLTQVMAHFEEM---GMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHR  776
          |||||||||||||||||||||    ||||||||:||.|.||||||||||||||||||||.
Db   622  HDDYDVR---YLAGCLGNISLLYMWGTDISSDQ------CSVLKQAIQQLPSIQ       666

Qy   777  STWSPTMVLFRWVPVTDAYWQILFSVLKVTRN------                      830
          .|||||||||||||||||||||||||||||:.
Db   667  ------VHQLYPDILSTYLINVARNWISNWFVSEVDLVPEALN                 702

Qy   831  TLRRPRCLLETLRLAGCGLTABDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLR  890
          |||||||||||||||.|||||:|||||||||||||||||||||||||||||||||||||.
Db   703  ------LGNGL------KDSKGQPRRNBELCSQDVSVPAPTQVNVNHYC---N        740

Qy   891  QPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDV-----------   937
          |||||||||||||||||||||||||||||||||||||||||.|||||
Db   741  APYHQVRFL---------TQAFILF--MINSLNIHPNGLDQVEKIWRLEDSEYLE     784

Qy   938  GVRLLCEGLRHPA---------------------CKLI-------RLGLD          959
          |||||||||.||                     ||||       |||||
Db   785  TPKLQCE---PASLIFKBQVTIKLQTHVVLDKETIRRCKLLHSKDGITWEBTKTKLEFS  840

Qy   960  QTTLS----------DEWRQELRALEQBKPQLLIFSRRKPSVMPTEGLD--TGEMS   1004
          |:||.          :|.|.||||||.||||||||.|||||||||||||  |||||
Db   841  DNFISFQTNYFSWWKVIFEFIEKTLAFLYKVLKSLLYQNAGDVVWKVCWNLDVATGB--  898

Qy  1005  NSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIABIAEBSSPEVVPVELLCVPB--  1061
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||.
Db   899  ---NELKENHFTQETWT--IPLSNDLILCLEK------HENEABVINI----IPSGKI   941

Qy  1062  -PASQ----------------GDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHFPVAGSY 1108
           ||.|                ||:|.|||.|.||||||.|||||:.|||.|.:||||||||||||
Db   942  IPANQLNNSYCCNKKFKVDKHS----TNE------VHLIAKAK------CGSF       978

Qy  1109  RWPNTGLCFVMREAVTVEIEPCVWDQ--FLGBINPQHSMWVAGPLLDIK           1155
          ||||||||||||.|.|.:|.||.|.   ||||||||||||||||||||:
Db   979  RW-DDDPCF------PLPITHAVSDQOPTQPSINP--TYIYTGPTTTIR          1019
```

RESULT 9
T31665
hypothetical protein COS1.4 - sea squirt (Ciona intestinalis)

```
C;Species: Ciona intestinalis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C;Accession: T31665
R;Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: T31665
A;Accession: Z21050
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1004 <BIR>
A;Cross-references: EMBL:Z80004; PIDN:CAB02586.1
C;Genetics:
A;Cross-references: EMBL:Z80004; PIDN:CAB02586.1
A;Introns: 30/1; 539/3; 593/3; 662/3; 728/3; 769/3; 817/2; 881/2; 960/1

Query Match             2.7%;  Score 202;  DB 2;  Length 1004;
Best Local Similarity   17.7%;  Pred. No. 7.9e-05;
  Matches 191;  Conservative 162;  Mismatches 373;  Indels 352;  Gaps 40;

Qy      244  AHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLQRPHPRSQDPLVKRSWPDY  303
                 :  :::   :                :  :       :   : : :   :
Db        9  AHAAVKIHQHP---------------------------------------------  30

Qy      304  VEENRGHLIEIRDLFGPGLDT-------------------QEPRIVILQGAAGIGKSTLAR  345
                :  :  : :  :                          :   ::::    : :
Db       31  ---TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTWVVGPHGGGKTTLLK  81

Qy      346  QVKEAWGRGQLYGDRFQHYFYFSCRELAQSKVVSLAELIGKDGTA----------TPAPIRQ  397
                :    :    :   : :: : :  :                         : ::
Db       82  MMVNQILKHELLPDI-EYIFIYAKDIDFNKEMTLLEFLITNSRVKVNYTEEESKALITF  140

Qy      398  ILSRPERLLFILDGVDEPGWVLQEFPSSELCLHWSQPQPADALLGSLLGKTILPEASFLIT  457
             :  :   :  : : :  :  ::: :   : : :  :  ::  : ::: : : :::
Db      141  LHNNPNVAIF-FDGLDEASLKELVGGYSICKLDEKSKPVD--IMKNLFNLALLPRAKIVT  198

Qy      458  ARTTALQNLIPLSEQARWEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALC  517
                  :  :  :  :: : : ::: :   :  :  ::     :   :  ::
Db      199  STPDEMFNLQHCYRPTSIFEVLGTLEBAXNNLGTQLCGEKYPAIK--KILDQOPNIAHLC  256

Qy      518  LVPFVVSWLACTCLM-----QQMKREKLTLTSKTTTTLCLHYLAQLAQAQPLGPQLRDLCS  573
             : :  : :                       :: : : : : ::::      :::::
Db      257  YLFINFILIVFCLLSNEGSDIKTMTQVLLFSMTRFVESSH-LKGEVPLDKVGAEMVKLAC  315

Qy      574  LAAEGIWQKKTLFSPDDLRKHGLDGAIISTF----------LKMGILQEHPIPLSYSFIH  623
             : :::::  : ::::::::: : :  : :          : : ::::   ::::::::
Db      316  LAYKGLQORKLVFEKTDFDDVKLADEMVINFHTYVDISSGIRIKILEGN--KRSY-FTH  372

Qy      624  LCFQEFFAAMSYVLEDEKGRGKGHSNCIIDLEKTLEAYGIHGLFGGASTTRFLIGLLSDEGE  683
                 :   :  :  :       :            :    :::  : : :  :
Db      373  HIWQBFYAAV-----------------------------YLMLFVSYREP  393

Qy      684  REMENIFHCRLSQGRNLMQWVPSLQLLL-----------QPHSLESLHCL-----YETRNKTLT  732
             :  :  :  :  :: :    :       ::::::       ::::::::
Db      394  EQLKTIFED-----------TQWSVVKFMFGICNPHAYKQLKIIFPATMIKDYEEKKKFLSS  445

Qy      733  QVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPV  792
                 :   :   : :  :   : : :     : ::   ::   :  : :  :: : :
Db      446  MVM---ESLSSAKSEDL-----IRRF-----GWLHEYNDDETSKKFKDCLPV  484

Qy      793  TDAYWQILFSVLKVTRNLKELDLSGGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAE  852
             :  :      :                   : :  :  :: :      :   :
Db      485  ------GLKMGVPKHLPEVKDLVYALKSFTKPHKLRLRSNWTTTTEVLETLLRGIHGTTTT  539

Qy      853  DCKDLAFGLRANQTLTELDLSFNVLTDDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQ  912
                :  :   :   :  :::: : :   :      : :  ::   ::   : :  :::
Db      540  ITRFVINNIEMKDSLMELLL---LHLDA-----MEKLRFDDVTNLSSYME  581

Qy      913  DLASVLSASPSLKELDL------QQNNLDDVGVRLLCEGIRHPACKLIRLGLDQTTLSDEMRQ  969
                :   :::  :: :     : : : : : : ::  :     : :: :  ::
Db      582  SLSNAINQRSNKIQLDLWHQOLNDDVVYLAGCLG-----NISRLNMSHTYLSSD--  632

Qy      970  ELRALEQEKPQLLIFSRRKPSVMPTEGLDTGEMSNSTSSLKRQLGSERAASHVAQANL  1029
             : :                 :: : : : :
Db      633  QCRVLKQAIEQL-----------------------PSI-----  647
```

```
RESULT 10

A55478
neuronal apoptosis inhibitory protein - human
N;Alternate names: NAIP
C;Species: Homo sapiens (man)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
C;Accession: A55478
R;Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird, S
d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; MacKenzie, A.
A;Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
A;Reference number: A55478; MUID:95112344; PMID:7813013
A;Accession: A55478
A;Molecule type: mRNA
A;Residues: 1-1232 <ROY>
A;Cross-references: GB:U19251
C;Genetics:
A;Gene: GDB:SMA®; SMA
A;Cross-references: GDB:120378; OMIM:600354; OMIM:253300
A;Map position: 5q12.2-5q13
A;Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prot
F;94-110/Domain: transmembrane #status predicted <TMM1>
F;470-477/Region: nucleotide-binding motif A (P-loop)
F;470-496/Domain: transmembrane #status predicted <TMM2>
F;476/Binding site: ATP (Lys) #status predicted
F;618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match             2.6%;  Score 198;  DB 2;  Length 1232;
Best Local Similarity   20.8%;  Pred. No. 0.0002;
  Matches 158;  Conservative 108;  Mismatches 282;  Indels 212;  Gaps 34;

Qy      292  QDPLVKRSWPDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAW  351
             :: :::        :   :  : : : :                :::    : :
Db      450  QEPLV-------------------LPEVFG-NLNS--------VMCVBGEAGSGKTVLLKKIIAFLW  487

Qy      352  GRGQL-YGDRFQHYFYFSCRELAQSK---VVSLAELIGKDGTATPAPIRQILSR-PERLL  406
                :  :   : :: : :  :          ::::    : :  : : :  :::::
Db      488  ASGCCPLLNFQLVFYLSLSSTRPDEGLASIICDQLLEKEGSVTEMCMRNIIQOLKNQVL  547

Qy      407  FILDGVDEPGWVLQEFPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNL  466
             :::::::::: : : : :   : : : :  ::  : ::: : : :::
Db      548  FLDDY--------KRIC---SIPQ----VIGKLIQKNHLSRTCLLIAVRTNRARDI  589

Qy      467  IPSLEQARWEVLGFSESSRKEYFYRYFTDERQAIRAFRL-VKSNKELWALCLVPMVWSWL  525
             : ::::::: : :::::::::::::::::::::::::: :::  :  : ::   :
Db      590  RRYLETI--LEIQAFPFYNTVCILRKLFSHNMTRLRKFMVYFGRKNQSLQKIQKTPLFVAA  647

Qy      526  ACTCLMQQ-------MKRKEKLTLTSKTTTTLCLHYLAQLAQAQPLGPQLRDLCS  573
             ::::: :         : : :  : :::::::::::::::::::::::::::
Db      648  ICAHWFQYPFDPSPDDVAVFKSYMRLSLENKAT--------ABILKATVSSCGE  694

Qy      574  LAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAM  633
             : :::::::::::::::::::::::::::::: : ::::::::::::::::  ::::::
Db      695  LALKGFFSCCFRFNDDLDABAGVDEDEDLTMCLMSKFTAQRLRPFYRFLSPAFQEFLAGM  754
```

```
Qy     1030  KLLDVSKIFPITAEIAEESSPEVVPVEILLCVPSPASQGDLHTKPLGTDDDFWGPTGPVATE  1089
             : : : :     : : :    :   :  :  :   :    :    : :   ::: :
Db      648  --QVHQLY----QVHQLY------PDILST-YLNVNRPIIRFDPNT------  674

Qy     1090  VVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVBIBFCVWDQFLGEINPQHSWMVAG  1149
             : : :  :  : : :    : : :  : : : :   :  : : :  :::
Db      675  --------SVYFVH----DQQPESSSKCWIGSRGGKLEVGGC----------  704

Qy     1150  PLLDIKAEPGAVE---AVHLPHFVALQGGGHVDTSLFQMAHFKEEGMLLEKPARVELH-HI  1205
             :::: :  :      ::   :  :  :  :  :   : : :   ::   : ::  :
Db      705  --ELVVPGALEKDVEIKLTASLSLESEFLETPILQC---ELASLTIKKQVTIKLQTHV  758

Qy     1206  VLENPSFSPLGVILLKMIHHNALRFIPVTISVVLLYRVVHPEEVTFHL---YLIPSD-CSI  1259
               : :   : :          : :   :    :  :::  :      :::  :
Db      759  VLDKET------------------IRRCKVTLVYTRV---TTTVHWGKGWLNHTDICSI  796
```

```
Qy    634  SYVLEDEKGRGKHSNC--IIDLEK--------TLEAYG------IHGL----FGASTTRFLLGL 677
                :| |||        ::|:           :||            ::  |   :| | ||
Db    755  RLIELLDSDRQEHQDLGLYIHLKQINSPMMTVSAYNNFLNYVSSLPSTKAGFKIVSHLIHL 814

Qy    678  LSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLHCLIYETRNKTFLTQVMAH 737
                :|:                                   :
Db    815  V--DNKESLENI----SENDDYLKHQPEISLQMQ------LLRGLWQICPQAYFSMVSEH 862

Qy    738  FEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGR--------QHRSTWSPTMV 784
                :|   :                                   |||||  |
Db    863  L--LVLALKIAYQSNTVAACSPF-----VLQFLQGRTLTLGALNLQYFFHPESLSLLRS 915

Qy    785  VLF-------------------RWVPVTDAYWQILFSVL-KVTRNLKELDLGSNSLS 821
           |||                   :|| ||||||||||||  |||||||| ||| | |
Db    916  IHFSIRGNKTSPRAHFSVLETCFDKSQVPFIDQDYASAFEPMNEWRNLAE------K 967

Qy    822  HSAVRSLCKTLRRPCLLETLRLAGCGLTAEDCKDLAFG---LRANQ-TLTELDLSFNVL 877
                :                                     |  :
Db    968  EDNVKSYMDMQRRA------------SPDLSTGYWKLSPKQYKIPCLEVDVNDi 1009

Qy    878  TDAGAKHLCQRLRQFSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDV 937
                                                                    :
Db   1010  DVVG--------QDMLEIILMTVFSASQRI-ELHLNHS----- 1037

Qy    938  GVRLLCEGLLRHPACKLIRLGLDQTTLSDEMRQELRALEQE 977
                 | :||  |||||||||||||||||||||| :  |: ||
Db   1038  --RGFIESIR-PALELSKASVTKCSIS---KLELSAAEQE 1071
```

RESULT 11
T42628
neuronal apoptosis inhibitory protein 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T42628
R;Yaraghi, Z.; Diez, E.; Gros, P.; MacKenzie, A.
Mamm. Genome 10, 761-763, 1999
A;Title: cDNA cloning and the 5'genomic organization of Naip2, a candidate gene for muri
A;Reference number: Z22179; MUID:99315342; PMID:10384056
A;Accession: T42628
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1447 <YAR>
A;Cross-references: EMBL:AF102871; NID:g3860228; PID:g3860229; PIDN:AAC73002.1
A;Genetics:
A;Gene: Naip2

```
Query Match          2.6%;   Score 196.5;   DB 2;   Length 1447;
Best Local Similarity   19.7%;   Pred. No. 0.00032;
Matches   172;   Conservative   138;   Mismatches   303;   Indels   261;   Gaps   43;

Qy    329  IVILGGAAGIGKSTLARQVKEAWGRGQL-YGDRFQIVFYFSC------RELAQSKVVSIA 381
              :|  |||:|||||  ||| :  :|   |  ||:|||| ||         :   :  ||
Db    509  VMCVGEAGSGKTTLKRIAFLWASGCCPLLNRFQLVFYLSLSSITPGQELAK--IICA 565

Qy    382  ELIGKDGTATAPIRQILSR-PERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALL 440
                |    |: |:|  ||   :|:  |:||  :  |   |:   :|       :   |
Db    566  QLLGAGGCISEVCLSSIIQQLQHQVLFLLD--DYSGLA 607

Qy    441  GSLLGKTILPEASFIITARTTALQNDIIPSLEQARWVEVLGFSESSRKEYFYRYFDERQA 500
              :|     :  |:|  || |  ||  :|:  :  | | |  |  |  |  |
Db    608  HTLITNQYLSRTCLLLIAVHTNKVRGIRPYLDTS--LEIKEFPYNTVSVLRKLFSHDIMR 665

Qy    501  IRAF-RLVKSNKELWALCLVPWVSWLACT------CLMQQMKRKEKLT 541
                :  |:|:|::||:|   :  :   | |        | : ||::|
Db    666  VRKFINYFGFHEELQGIHKPLFVAAVCTDWFKNPSDQPFQDVALFKAYMQYLSIKHKGA 725

Qy    542  LTSKTTTLCLHYLAQALAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLD--GAI 600
           ::|                                                     ||:
Db    726  -----AKPLQATVSSCQLALTGLFSSCFEFNSDNLAEAGVDEDEE 766

Qy    601  ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAA------MSVVLEDEKGRGKHSNCIIDLE 654
              :||                                  : |||||||||||||||||||||
```

```
Db    767  LTTCLMSKFTAQRLRPV-YRFLGPLFQEFFLAAVRLTELILSDSDRQEDQDLGLY--YLRQIN 823
                                :|     :|:|||  | |                :   |::|
Qy    655  KTLEAYGIHGLF-----------------GASTTRFLLGLSDEGEREMENIFHCRLSQGRNLMQW 703

Db    824  SPLKAMSIYHTFLKYVSSHPSSKAAPTVSHLLQLVDEKE-SLEN----MSENEDYMKL 877
                |:                                    ||:|
Qy    704  VPSLQLLLQPHSLESLHCLIYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRH 763

Db    878  HFEALLW------IECLRGLWQLSPESFSLPISEN---------LLRICLNFAHE 917
              : |||         :|:|:|||   :    :: :            ||::|  :|
Qy    764  VKK------------LQLIEGR---------QHRSTWS-PTMVVLFR---------WV----- 790

Db    918  SNTVAACSPVILQFLRGRTLDLKVLSLQYFWDHPETLLLLKSIKISLNGNNMVQRIDPSL 977
                                || ||  |   ||  ||
Qy    791  ------------PVTDAYWQILFS-VLKVTRNLKEL--------DLSGN--SLSHSAVKSL 828

Db    978  IEKSFEKVQPPFIDQDYAIAFQPINEVQKNLSEKKHIIKKYEDMKHQIPLNISTGYWKLS 1037
                      || ||||| ||  |
Qy    829  CKTLRRPCLLETLRLAGCGLTAEDCKDLAFGLRANQTLTEL-------DLSFNVLTD 879

Db   1038  FKPYKIPK--LE-VQVTNTG--------------PADQALLQVLMEVFSASQSIEFRLSDS 1081
                                                : :|  |: | |:|
Qy    880  AGAKHLCQRLRQFSCKLQRLQLVSCGLTS-DCCQDLASVLSASPSLKELDLQQ-NNLDD 936

Db   1082  SG---FLESIR-PALELSKASVTKCSMSRLEISREDQKLLLTLPTLQSLEVSETNQLPDQ 1137
                   ||||| |||||||||||||| |                 |
Qy    937  ------------VGVRLLC-------------EGLRHPACKLIRLGLDQTTLSDEMRQEL 971

Db   1138  LFHNLHKFLGLKELCVRLDSKPDVLSVLPGEPFNLHH----MEKLSIRTSTESD--LSKL 1191
               :                    |:        |  :|
Qy    972  RALEQEKPQLLIF--------SRRKP--SVMTPTEGLDTGEMSN---------------- 1005

Db   1192  VRLIQNSPNLHVFHLKCNFLSNCEPLMTVLASCKKLREIEFSGRCFEAMTFVNILPNFVF 1251
                                                 :
Qy   1006  -STSSLKRQLGSERAASHVAQANLKLLDVSKIF 1038

Db   1252  LKILNLRDQQFPDKETSEKFRAQALGSLRNLEKLF 1285
```

RESULT 12
T52063
ran GTPase-activating protein [imported] - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52063
R;Pay, A.; Nick, P.; Nagy, F.
submitted to the EMBL Data Library, December 1999
A;Reference number: Z25929
A;Accession: T52063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-533 <PAY>
A;Cross-references: EMBL:AF215731; PIDN:AAF19528.1
A;Experimental source: subspecies x varia; cultivar A2

```
Query Match          2.4%;   Score 180;   DB 2;   Length 533;
Best Local Similarity   21.1%;   Pred. No. 0.0008;
Matches   110;   Conservative   76;   Mismatches   188;   Indels   148;   Gaps   16;

Qy    544  SKTTTTLCLHYLAQALAQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIST 603
                                                  :
Db     89  AKESSKLMLEVLKRGPRGKEENGEL-----ISEKGDAAVEVTVFDISGGRRAFIDGQEASE 143

Qy    604  FLKMGILQEHPIPLSYSFIHLCF--QEFFAAMSYVLBDEKGRGKHSNCIIDLEKTL-- 657
                                                  :
Db    144  LLK-----PLMGPNSFTKICFSNRSFGLDAAHVEPMLISIKDQLKEVDLSDFIAGRP 196

Qy    658  --EAYGIHGLFGASTTRFLL----------------GILSDEGEREMENIFHCRLSQ 696
                                                   :|
Db    197  EAEAIEVMNIFSSALERAVLRYLNLSNNAMGEKGVRAFRALLKSQNDLEELYLMNDGISE 256

Qy    697  --GRNLMQWVPVSLQLLLQPH-------------SLESLHCLIYETRNKTFLT 732
                  |:                           ||:|||||||||||||||
```

RESULT 13
JC7765
mitotic spindle associated protein, MAP126 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: JC7765
R;Chang, M.S.; Huang, C.J.; Chen, M.L.; Chen, S.T.; Fan, C.C.; Chu, J.M.; Lin, W.C.; Yar
Biochem. Biophys. Res. Commun. 287, 116-121,2001
A;Title: Cloning and characterization of hMap126, a new member of mitotic spindle-associ
A;Reference number: JC7765; PMID:11549262
A;Contents: Testis
A;Accession: JC7765
A;Molecule type: mRNA
A;Residues: 1-1120 <CHA>
A;Cross-references: GB:AF345347
A;Comment: This protein, as a nuclear protein and a new member of kinesin family, which
and dynamic regulation of mitotic spindles and may provide an outward force to push the
C;Genetics:
A;Genetics:
A;Map position: 17q11-12

Query Match   2.2%; Score 166.5; DB 2; Length 1120;
Best Local Similarity   18.8%; Pred. No. 0.02;
Matches 217; Conservative 138; Mismatches 391; Indels 411; Gaps 50;

Qy   2 WRVKKLSLSLSPSPQTGKKSMRTLPRELTLQPGALTTSGK-RSPACSSLTPSLCKLGLQE   60
Db   61 GSNNSSPVDFVNNKRTDLSSEHFSHSSKWLFETCQHEBSDEQPLDPIPQISSTRKTISEBAVD 120
Qy 201 APGTQWPLD------ETSGIYYTE---IREREREKSEKGRPWAAVVGTPPQAHTSLQ 249
Db 250 PHHHPWEPSVRESLCSTWPWKNEDNRDNQKFTQLLLLQRPHPRSQDPLVKRSWPDVVEENRG 309
Qy 310 HLIEIRLFQPGLDTQ-EPRIVILQGAAGIGKSTLARQVKEAWGRQOLYGDRFQHVFYFS 368
Db 121 PLG----------------NYMVKTIVLVPSPLGQQODM--- 143
Qy 144 -----IFEARLDTMAETNSISILNGP--LRTDDLVREEV-----APCMGDRF-- 182
Db 369 CRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDPGWVLQEPSSELC- 427
Qy 183 ------SEVAAVSE--------------KPIFQ- 211
Db 428 --LHWSQPQPADALLGSLLGKTILPEASFLITARTTAL-QNLIPSLEQA----RWVEV 478
Qy 212 EQLHCSKES----------------LSSRIEAVREDLVPSESNAFLPSSVLW--- 247
Db 479 LGFSSSSRKEYFYRYFTDEROAI-------RAFR-----LVKSNKELWALC 517

Qy 148 WR--EISASLLYQALPSSPDHESPSQESPNAPTSTAVLGSWGGSPPQPSLAPR-----EQE 200

RESULT 14
A45841
T-complex-associated-testes-expressed-1 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
C;Accession: A45841
R;Sarvetnick, N.; Tsai, J.Y.; Fox, H.; Pilder, S.H.; Silver, L.
Immunogenetics 31, 283-284, 1990
A;Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
A;Reference number: A45841
A;Accession: A45841
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-506 <SAR>
A;Cross-references: GB:M28821; NID:g201909; PIDN:AAA40406.1; PID:g201910

Query Match   2.1%; Score 159.5; DB 2; Length 506;

Db 248 LSPSTALAADFRVRNHVDPEEEIVEHGAMEREMRFPTHPKESETEDQALVSSVEDIISTC 307
Qy 734 VMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQ------LIEGRQHS 777
Db 308 LTPNLVEMESQEAPGPAVEDVGRILGSDTESWMSPLA 353
Qy 778 TWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPC 837
Db 518 LVP-----------------WVSWLACTCLMQQMRKEKLITLSKTTTT 549
Qy 552 KLRAKLQSLKAEREEARHREEMALRGKDAARIVLEAFCAHASQRISQLEQDLASMREFRG 611
Db 734 ------VMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQ 777
Qy 612 LLKDAQTQL---VGLHAKQESELVQQTVSLTSTLQQ-DWRSMQLDYTTWTAL--LSRSRQ 664
Db 838 LLETLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSPNVLTDAG--AKHLCQRLRQPSCK 895
Qy 665 LITEKLT---------VKSQQALQERDVAIEEKQEVSRVLBQVSAQLEECKGQ 707
Db 896 LQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNN-LDDVGVR---LLCE----- 944
Qy 708 TEQLELENIRLATDLRAQLQILANMDSQLKELQSQHTHCAQDLAMKDELLCQLTGSNEEQ 767
Db 945 -----GLRHPACKLIRLGLDQTTLSDEMRQELRALE------- 977
Qy 768 AAQCVKREMALKHMQAELQQ-----QQAVLAKEVRDLKETLEFPADQENQVAHLELGQVECQL 824
Db 978 KPQLLIFSRR------KPSVWMTPTPEGLDTGEMSNSTSSLKRQRLGSSERAASHVAQANLK 1030
Qy 825 KTTLEVLRERSLQCENLKDTVENLTAKLASTIADNQEQDLEKTRQYSQKLGLITEQ--- 880
Db 1031 LLDVSKIFPIAEIAEESSPEVVPVEILLCVPSPASQGDLIHTKPLGTDDDFWGPTGPVATEV 1090
Qy 881 -LQSLTLFLQTKLKERTEQETLLLSTACPPT-----QEHPLPNDRTF---LGSILTAV 929
Db 1091 VDKEKNLYRVHFPVAGS 1107
Qy 930 ADEEPESTPV--PLLGS 944

Query Match   2.1%; Score 159.5; DB 2; Length 506;

```
RESULT 15
T47442
disease resistance protein homolog - Arabidopsis thaliana
N;Alternate names: protein T18B22.70
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47442
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A;Reference number: 224467
A;Accession: T47442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1199 <JOR>
A;Cross-references: EMBL;AL138652
A;Experimental source: cultivar Columbia; BAC clone T18B22
C;Genetic:
A;Genetics:
A;Map position: 3
A;Introns: 57/1; 238/2; 613/3; 713/3
A;Note: T18B22.70

Query Match     2.1%;  Score 155.5;  DB 2;  Length 1199;
Best Local Similarity  19.4%;  Pred. No. 0.12;
Matches 222;  Conservative 141;  Mismatches 361;  Indels 421;  Gaps 57;
```

```
Best Local Similarity  20.5%;  Conservative  77;  Pred. No. 0.016;
Matches 117;                   Indels 183;  Gaps 25;

Qy   501 IRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKILTLSKTTTTLCLHYLAQALQ 560
Db    55 VRRMRRIISEDAEWSLAIVPLL----------------TELCIQHIVANFQ 89

Qy   561 AQPLGPQLRDLCSLAAEGIWQKKTLFS-PDDLRKHGLDGAIISTFLKMGILLQEHPIPLSY 619
Db    90 NNPI----LKQLPLB--HQKKVLSNLPPEL- 117

Qy   620 SFIHLCPQEFFAAMSYVLEDEKRGRGKISNCIIDLEKYLEAYGI------HGLFGASTTRFL 674
Db   118 T-------ANIIDDE--NYWHRCCI----------KRWSVCHVSRHG--GSWKRMPF 153

Qy   675 LGLLSDEGEREMENIFH------CRLSQGR-NLMQWVPSLQLILQPHS 715
Db   154 ------ERHLBNLLFIPGTTDPNVILDLLPLCRNYVRIHVDQFLPVRM--PTP 202

Qy   716 L--------ESLHCLYETRNKTFLTQVMAHFEEM------GMCVETDM 749
Db   203 LQGEEQOSDSGSGSBEGEGSSPEKDHYQLQ?-----LVGGLKHLBELDLVYGVKDQCGMNPEWNL 257

Qy   750 ELLVCTFPCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWV--PVTDAYWQILFSVLKVT 807
Db   258 FLFTYRDCYSLAAATIKACHTLK-----------IPFKLTRSKVDDDKARILIRSLLDH 303

Qy   808 RNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTL 867
Db   304 PALELELDLSHLIGDRGARAAAKLLSHSR--LRVLNLANNQLQAPGAQSLAHALANNTNL 361

Qy   868 TELDLSPNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKEL 927
Db   362 VFLNLRNCIEDEGGQAIAHALETNKC-LSVLHLGGNKLSEPTATLLSQMLTVNTLVSL 420

Qy   928 DLQQNNLDDGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEKPQLL--IFS 985
Db   421 NLSCNHLGQDGGKQLLEGISD---------NKTLLEFDLR--LSDVSQESEYLIGQVLH 468

Qy   986 RRKPSVMTPTEGLDTGEMSNSTSSSLKRQRLG 1016
Db   469 ANREAARQRT--LNPGHFSSPTNNCTENSVG 497
```

```
Qy   119 PWIHELPAGCTQGSERRVLRQLPDTSGRRWRE-------ISASLILYQALPSSPDHBSPSQBS 173
Db    92 -WKHHV-------FPSFHGADVRKTILSHILESFRRKGIDPFIDNNIERSKS 135

Qy   174 PNAPTSTAVLGSWGSPPQPSLAPREQEAPGTQWPLDETS---------GIYY-- 216
Db   136 IGHELKEAIKGS-------KIAIVLLSKNYASSWCLDELAEIMKCRELLGQIVWTIFYBVD 190

Qy   217 -TEIRERERE----------KSEKGRPPWAAVVGTPPQAHTSLQPHHHPWEPSVRESLC---S 265
Db   191 PTDIKKQCTGBFGKAFTKTCKGK---------TKEYVERWRKALEDVATIAGEH 234

Qy   266 TWPRKNE-DFNQKF---TQLLLQRPHPRSQODPLVKRSWPDYVEBNRGHLIEIRDLFGPG 321
Db   235 SRNWRNRADMIERKIATDVSNMLNSFTPSRDFPDGLVGM---------RAHMDMLBQLL--R 283

Qy   322 LDTQBPRIVILQGAAGIGKSTLAR------QVKEAWGRGQLYGDRPQ-------HVFY-- 366
Db   284 LDLDEVRMIGIWGPPGIGKTTIARFLPNQV--------SDRFQLSAIIVNIRGIYPRP 333

Qy   367 ----FSCRBLAQSKVVSLAELLG-KDGTATPAPIRQILSRPBRLLFILDGVDE----- 414
Db   334 CPDEYSAQLQLQNQM--LSQMINHKDIMISHLGVAQBRLRDKKVFLVLLDEVDQLGQLDAL 391

Qy   415 ------PG----------WTLQBPSSE------LCLH-WSQPQPAD-- 437
Db   392 AKBTRWFGPGSRIIITTBDLGVLKAHGINHVVKVKYPSNDEAFQIFCWNAFGQKQPHEGF 451

Qy   438 -------ALLGSL-LGKTIL------PEASFLITARTTALQNLIPSLEQARWVEVL 479
Db   452 DEIAWEWALAGELPLGLKVLGSALRGKSKPEWERTLPRLKTSLDGNIIGSIIQ----- 504

Qy   480 GFSESSRKEYFRVFTDERQATRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKREK 539
Db   505 ----------LLYIACLPNYESTTKVEE 535

Qy   540 LILTLSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGA 599
Db   536 VLANKFLDVKQGLHVLAQ------KSLISIDENSLYG-DTI 569

Qy   600 IISTFLKMGILQLQEHPIPLSYSFIHLCPQBFFAAMSYVLBDEKGR-----GKHSNCIIDLEK 655
Db   570 NMHTLLRQ-------FGRETSRKQFVVHGFTKRQLLVGERDICBVLSDD 611

Qy   656 TLEAYGIHG-----LFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLL 711
Db   612 TIDSRRFIGITFPDLFG--TQDYL--NISEKALERMNDFEFVRI----NALIFTERLQLAL 663

Qy   712 QP------HS--LBSLH---------CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIK 759
Db   664 QDLICHSPKIRSLKWYSYQNICLPSTNPEFLVELHMSFSKL----- 705

Qy   760 FSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNS 819
Db   706 --------RKLWEGT- 726

Qy   820 LSHSAVKSLCKTLRRPRCLLETLRLACGLTAEDCKDLAFGLRANQTLTELDLSPNVLTD 879
Db   727 EDLKELPNLSTATN----LEELKIRDCS---------SLVELPSSIEKLTS 764

Qy   880 AGAKHL--CQRLRQ-PS-----CKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKELDLQQN 932
Db   765 LQRLYLQRCSSLVBLPSFGNATKLEELYLENC--------SSLBKLPPSINANNLQOL 814

Qy   933 NLDDGVGRLLCEGLRHPACKLIRLGLD------QTTLSDEMRQELRALEQBKPQLLIFSR 986
Db   815 SLINCSRVVBLPAIEN-ATNLQKLDLGNCSSLIEPLSIGTATNLKELNISGCSSLV-- 870

Qy   987 RRKPSVMTPTEGLDTGEMSNSTSSSLKRQRLGSRAASHVAQANLKLLDVSKIFPIAEIAEE 1046
Db   871 KLPSSIGDITNLKBPDLSNCSNLVELP---------ININKFLDTLNLAGCSQL--K 917

Qy   63 EQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSF----PYSPSEPHLGSPSQPTSTAVLM 118
Db   43 DESSWK---HPW------SLCVNVAAAFTKFRFQQDNKYTKSSA-LSLPSPPTSVSRI- 91

Qy  1047 SSPBV 1051
```

Db | |||:
      918 SFPEI 922

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 09:40:41 ; Search time 21.9982 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWGRLACYLEFLKKEE........HLIMELWEKGSKKGLLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	7502	99.6	1473	1	NAL1_HUMAN	Q9C000	homo sapien
2	1382.5	18.4	1062	1	PYA7_HUMAN	Q59046	homo sapien
3	1279	17.0	1034	1	CIS1_HUMAN	Q96P20	homo sapien
4	1254.5	16.7	1033	1	CIS1_MOUSE	Q8r4b8	mus musculu
5	1037	13.8	994	1	NAL4_HUMAN	Q96mn2	homo sapien
6	983	13.0	1200	1	MATE_HUMAN	P59047	homo sapien
7	948	12.6	1062	1	NAL2_HUMAN	Q9nx02	homo sapien
8	908.5	12.1	980	1	PYA3_HUMAN	Q8wx94	homo sapien
9	901.5	12.0	892	1	PYA5_HUMAN	P59044	homo sapien
10	895	11.9	854	1	PYA5_RAT	Q63035	rattus norv
11	882	11.7	1111	1	MATE_MOUSE	Q9r1m5	mus musculu
12	868	11.5	843	1	PYA5_MOUSE	Q91w82	mus musculu
13	711	9.4	1033	1	PYA6_HUMAN	P59045	homo sapien
14	668	8.9	431	1	CAR8_HUMAN	Q9y2g2	homo sapien
15	393.5	5.2	1040	1	CARF_HUMAN	Q9hc29	homo sapien
16	392	5.2	1020	1	CARF_MOUSE	Q8k3z0	mus musculu
17	388.5	5.2	953	1	CAR4_HUMAN	Q9y239	homo sapien
18	361.5	4.8	456	1	RINI_HUMAN	P29315	rattus norv
19	358	4.8	460	1	RINI_RAT	P13489	homo sapien
20	354.5	4.7	456	1	RINI_PIG	P10775	sus scrofa
21	351	4.7	953	1	CAR4_MOUSE	Q8bhb0	homo sapien
22	345	4.6	1130	1	C2TA_HUMAN	P33076	homo sapien
23	318	4.2	1155	1	C2TA_MOUSE	P79621	mus musculu
24	253.5	3.4	193	1	ASC_MOUSE	Q9epb4	mus musculu
25	250.5	3.3	195	1	ASC_HUMAN	Q9ulz3	homo sapien
26	215	2.9	1024	1	CARC_HUMAN	Q9npp4	homo sapien
27	212.5	2.8	1403	1	CARC_HUMAN	Q9r016	homo sapien
28	200	2.7	1403	1	BIRE_MOUSE	Q9jib6	mus musculu
29	199	2.6	1402	1	BIRG_MOUSE	Q9jib3	mus musculu
30	198	2.6	1403	1	BIR1_HUMAN	Q13075	homo sapien
31	196.5	2.6	1447	1	BIRB_MOUSE	Q9quk4	mus musculu
32	196	2.6	1447	1	BIRA_MOUSE	Q9qwk5	mus musculu
33	161	2.1	483	1	YAX2_HUMAN	O95522	homo sapien
34	158	2.1	5065	1	EPPL_HUMAN	P58107	homo sapien
35	152	2.0	3511	1	MY15_MOUSE	Q9qzz4	mus musculu
36	147.5	2.0	793	1	ST5A_MOUSE	P42230	mus musculu
37	146.5	1.9	1524	1	Y133_HUMAN	O14146	homo sapien
38	145.5	1.9	793	1	ST5A_RAT	Q62771	rattus norv
39	144	1.9	2300	1	CYAA_NEUCR	O01631	neurospora
40	142.5	1.9	4303	1	NPH4_HUMAN	P98161	homo sapien
41	142	1.9	1426	1	PKD1_HUMAN	O75161	homo sapien
42	142	1.9	1839	1	CYAA_SACKL	P23466	saccharomyc
43	141.5	1.9	1021	1	PSKR_DAUCA	O81pb4	daucus caro
44	140	1.9	794	1	ST5A_BOVIN	Q95115	bos taurus
45	138.5	1.8	794	1	ST5A_HUMAN	P42229	homo sapien

ALIGNMENTS

RESULT 1
NAL1_HUMAN
ID   NAL1_HUMAN   STANDARD;   PRT;   1473 AA.
AC   Q9C000; Q9BZZ8; Q9HAV8; Q9UFT4; Q9Y2E0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   NACHT-, LRR- and PYD-containing protein 1 (Death effector filament-
DE   forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
DE   caspase recruitment domain) (Caspase recruitment domain protein 7).
GN   NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=21148093; PubMed=11250163;
RA   Martinon F., Hofmann K., Tschopp J.;
RT   "The pyrin domain: a possible member of the death domain-fold family
RT   implicated in apoptosis and inflammation.";
RL   Curr. Biol. 11:R118-R120(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=21169419; PubMed=11270363;
RA   Bertin J., DiStefano P.S.;
RT   "The PYRIN domain: a novel motif found in apoptosis and inflammation
RT   proteins.";
RL   Cell Death Differ. 7:1273-1274(2000).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX   MEDLINE=21153743; PubMed=11076957;
RA   Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
RA   Shi M.M., Vincenz C., Ward P.A.;
RT   "Molecular cloning and characterization of DEFCAP-L and -S, two
RT   isoforms of a novel member of the mammalian Ced-4 family of apoptosis
RT   proteins.";
RL   J. Biol. Chem. 276:9230-9238(2001).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.
RC   TISSUE=T-cell;
RX   MEDLINE=21153744; PubMed=11111115;
RA   Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
RA   Godzik A., Reed J.C.;
RT   "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
RT   c-dependent caspase activation and apoptosis.";
RL   J. Biol. Chem. 276:9239-9245(2001).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code

RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [6]
RP   SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
RC   TISSUE=Uterus;
RA   Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Able to form cytoplasmic structures termed death
CC   effector filaments. Enhances APAF1 and cytochrome c-dependent
CC   activation of pro-caspase-9 and consecutive apoptosis. Seems to
CC   bind ATP.
CC   -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
CC   and with APAF1 in a cytochrome c-inducible way leading to the
CC   formation of an apoptosome. This interaction may be ATP-dependent.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=4;
CC   Name=1; Synonyms=NAC beta, DEFCAP-L;
CC   IsoId=Q9C000-1; Sequence=Displayed;
CC   Name=2; Synonyms=NAC alpha, DEFCAP-S;
CC   IsoId=Q9C000-2; Sequence=VSP_004327;
CC   Name=3; Synonyms=NAC gamma;
CC   IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
CC   Name=4; Synonyms=NAC delta;
CC   IsoId=Q9C000-4; Sequence=VSP_004326;
CC   -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
CC   expressed in peripheral blood leukocytes, chronic myelogenous
CC   leukemia cell line K-562 followed by thymus, spleen and heart.
CC   Also detected in lung, placenta, small intestine, colon, kidney,
CC   liver and muscle.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF298548; AAG15254.1; -.
DR   EMBL; AF310105; AAG30288.1; -.
DR   EMBL; AF229059; AAK00748.1; -.
DR   EMBL; AF229060; AAK00749.1; -.
DR   EMBL; AF229061; AAK00750.1; -.
DR   EMBL; AF229062; AAK00751.1; -.
DR   EMBL; AB023143; BAA76770.1; -.
DR   EMBL; AL117470; CAB55945.1; -.
DR   PIR; T17255; T17255.
DR   HSSP; P13489; 1A4Y.
DR   MIM; 606636; -.
DR   GO; GO:0005622; C:intracellular; IC.
DR   GO; GO:0016505; F:apoptosis activator activity; NAS.
DR   GO; GO:0008656; F:caspase activator activity; NAS.
DR   GO; GO:0019899; F:enzyme binding activity; IPI.
DR   GO; GO:0006919; P:caspase activation; NAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00560; LRR; 2.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PRINTS; PR00364; DISEASERSIST.
DR   PROSITE; PS50209; CARD; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;

KW   Alternative splicing.
FT   DOMAIN   1   92   DAPIN.
FT   DOMAIN   328   637   NACHT.
FT   REPEAT   704   725   LRR 1.
FT   REPEAT   807   830   LRR 2.
FT   REPEAT   864   887   LRR 3.
FT   REPEAT   921   944   LRR 4.
FT   REPEAT   950   973   LRR 5.
FT   REPEAT   1199   1215   LRR 6.
FT   REPEAT   1216   1236   LRR 7.
FT   DOMAIN   1374   1463   CARD.
FT   NP_BIND   334   341   ATP (POTENTIAL).
FT   VARSPLIC   958   987   Missing (in isoform 3 and isoform 4).
FT   /FTId=VSP_004326.
FT   VARSPLIC   1262   1305   Missing (in isoform 2 and isoform 3).
FT   /FTId=VSP_004327.
FT   MUTAGEN   340   340   K->L: NO EFFECT.
FT   MUTAGEN   340   340   K->S: NO EFFECT.
FT   CONFLICT   155   155   L -> H (IN REF. 1).
FT   CONFLICT   246   246   L -> S (IN REF. 1).
FT   CONFLICT   782   782   T -> S (IN REF. 1).
FT   CONFLICT   878   878   T -> M (IN REF. 1).
FT   CONFLICT   995   995   T -> I (IN REF. 1).
FT   CONFLICT   1119   1119   M -> V (IN REF. 1).
FT   CONFLICT   1184   1184   M -> L (IN REF. 1 AND 6).
FT   CONFLICT   1241   1241   V -> M (IN REF. 1).
FT   CONFLICT   1366   1366   R -> C (IN REF. 1).
SQ   SEQUENCE   1473 AA; 165865 MW; 438F0DCE45C25E2D CRC64;

Query Match   99.6%; Score 7502; DB 1; Length 1473;
Best Local Similarity   97.0%;
Matches 1429; Conservative   0; Mismatches   0; Indels 44; Gaps   1;

Qy   1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWBQMGLRSLCAQAOEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWBQMGLRSLCAQAOEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQBSSPNAPTST 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQBSSPNAPTST 180

Qy   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBKSEKGRPPWAAVGT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERBKSEKGRPPWAAVGT 240

Qy   241 PPQAHTSLQPHHHPWEPSVRBSLCGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 PPQAHTSLQPHHHPWEPSVRBSLCGSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy   301 PDIYEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PDIYEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy   481 FSESSRKEYFYRYFTDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAPRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy   541 TLTSKTTTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTTLCLHYLAQALQAQPLGPQLRDLCSLAABGIWQKKTLFSPDDLRKHGLDGAI 600

```
Qy   601  ISTFLKMGILQEHPIPLSYSFIHLCPQBFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  ISTFLKMGILQEHPIPLSYSFIHLCPQBFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660

Qy   661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH  720

Qy   721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780

Qy   781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVTKSLCKTLRRPRCLLE  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVTKSLCKTLRRPRCLLE  840

Qy   841  TLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ  900

Qy   901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960

Qy   961  TTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLSLKRQRLGSSERA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTLSDEMRQELRALEOBKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLSLKRQRLGSERA  1020

Qy  1021  ASHVAQANLKLLDVSKIPPIAEIAEBESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ASHVAQANLKLLDVSKIPPIAEIAEESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080

Qy  1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMRBAVTVEIEFCVWDQPLGBIN  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMRBAVTVEIEFCVWDQPLGBIN  1140

Qy  1141  PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDDTSLFQWMAHFKEEGMLLEKPARV  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  PQHSWMVAGPLLDIKAEPGAVEAAVHLPHFVALQGGHVDDTSLFQWMAHFKEEGMLLEKPARV  1200

Qy  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy  1261  -----------------------------KELELCYRSPGEDQLF  1276
                                       |||||||||||||||||
Db  1261  KAIDDLEMKFQFVRIHKPPLPLTPLYMGCRYTVSGSGSGMLEILPKELELCYRSPGEDQLF  1320

Qy  1277  SEFYVGHLGSSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  SEFYVGHLGSSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH  1380

Qy  1337  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLPSLSQSWDRKC  1396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSGMRKLPSLSQSWDRKC  1440

Qy  1397  KDGLIYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
          ||||||||||||||||||||||||||||||||||
Db  1441  KDGLIYQALKETHPHLIMELWEKGSKKGLLPLSS  1473
```

RESULT 2
PYA7_HUMAN

```
ID   PYA7_HUMAN     STANDARD;      PRT;   1062 AA.
AC   P59046;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYRIN-containing APAF1-like protein 7 (Monarch-1).
GN   PYPAF7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC   TISSUE=Lymphoma;
RA   Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Leukocyte;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunarante P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```
-!- FUNCTION: May mediate activation of CASP1 via ASC and promote
    activation of NF-kappa-B via IKK.
-!- SUBUNIT: Binds to ASC with its DAPIN domain.
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- ALTERNATIVE PRODUCTS:
    Event=Alternative splicing; Named isoforms=3;
    Name=1; Synonyms=I;
      IsoId=P59046-1; Sequence=Displayed;
    Name=2; Synonyms=II;
      IsoId=P59046-2; Sequence=VSP_005523;
    Name=3; Synonyms=III;
      IsoId=P59046-3; Sequence=VSP_005524;
-!- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
    predominantly in eosinophils and granulocytes, and at lower levels
    in monocytes.
-!- SIMILARITY: Contains 1 DAPIN domain.
-!- SIMILARITY: Contains 1 NACHT domain.
-!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
```

```
DR   EMBL; AY095146; AAM18227.1; -.
DR   EMBL; AY116204; AAM75142.1; -.
DR   EMBL; AY116205; AAM75143.1; -.
DR   EMBL; AY116206; AAM75144.1; -.
DR   EMBL; BC028069; AAH28069.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
```

```
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN.
DR   Pfam; PF00560; LRR; 2.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00368; LRR_RI; 11.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS00837; NACHT; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
FT   DOMAIN        1     95       DAPIN.
FT   DOMAIN      211    528       NACHT.
FT   REPEAT      713    736       LRR 1.
FT   REPEAT      742    765       LRR 1.
FT   REPEAT      770    792       LRR 2.
FT   REPEAT      799    822       LRR 3.
FT   REPEAT      827    850       LRR 4.
FT   REPEAT      884    907       LRR 5.
FT   REPEAT      941    969       LRR 6.
FT   REPEAT      998   1021       LRR 7.
FT   REPEAT     1021              LRR 8.
FT   NP_BIND     217    224       ATP (POTENTIAL).
FT   VARSPLIC    863    974       Missing (in isoform 3).
FT                                /FTId=VSP_005523.
FT   VARSPLIC    977   1032       Missing (in isoform 2).
FT                                /FTId=VSP_005524.
FT   CONFLICT    692    692       MISSING (IN REP. 1 AND 3).
FT   CONFLICT    977   1032       MISSING (IN VSP_005524).
SQ   SEQUENCE 1062 AA; 120328 MW; 0AB1C87F116497P CRC64;

Query Match             18.4%;  Score 1382.5;  DB 1;  Length 1062;
Best Local Similarity   34.4%;  Pred. No. 7.4e-81;
Matches  371;  Conservative  143;  Mismatches  373;  Indels  193;  Gaps  24;

Qy      8 RLACYLFPLKKEELKEPQLLLANKAHSRSSSGETP-AQPEKTSGMEVASYLVAQYGEQRA 66
            ||||:|| |||:||  ||||||  |:|||   : |:||||
Db     12 RLSTYLEELEAVELKKFKLYLGTA--TELGEGKIPWGSMEKAGPLEMAQLLITHFGPEEA 69

Qy     67 WDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSBSEPHLGSPSQ-PTSTAVLMPWIHELP 125
Db     70 WRLALSTFERINRKDLWERGQREDLVRDTPPGGPSS--LGNQSTCLLEVSLVTP--RKDP 125

Qy    126 AGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTSTAVLGS 185
Db    126 QEBTYRDYVRRKFRLMEDNRARLGECVNLSHRYTRLLLVKEHSNPMQVQ----------- 173

Qy    186 WGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAAVGTPPQAH 245
Db    174 --------------------------------------------------------- 173

Qy    246 TSLQPHHPWEPSVRESLCSTWPKNEDNQKFTQLLLQRPHPRSQDPLVKRSWPDYVE 305
Db    174 ------------QOLLDTGRGHARTV------------------------- 187

Qy    306 ENRGHL---IEIRDLFGPGLDTQF-PRIVILQGGAAGIGKSTLARQVKEAWGRGQLYGDRP 361
Db    188 ---GHQASPIKIFTLFEPDEERPEBPPRTVVMQGAAGIGKSMLAHKVMLDWADGKLFQGRF 244

Qy    362 QHVFYFSCRELAQSKV-VSLAELIGKDGTATPAPIRQIISRPERLLFILDGVDEPGWVLQ 420
Db    245 DYLFYINCREMNQSATECSMQDLIFSCWPEPSAPLQELIRVPERLLPIIDGPDELKPSFH 304

Qy    421 EPSSELCLHWSQPOPADALLGSLLGKTILPEASFLITARTTAIQNLIPSLEQARWVEVLG 480
Db    305 DPQGPWCLCWBEBKRPTELLLNSLIRKLLPELSLLITTRPTALEKLHRLLEHPRHVEILG 364

Qy    481 FSESSRKEYFYKFYRYFTDERQAIRAFRLVKSNKELWALCLLVPWVSWLACTCLMQQMRKEKQL 540
Db    365 FSEAERKEYFYKFYHNAEQAGQVNYVRDNEPLPTMCFVPLVCWVVCTCLQQOLEGGGLL 424

Qy    541 TLTSKTTTTLCLHYLLAQALQAQPLGPQL------RDLCSLAAEGIWQKKTLFSPDDLRKH 594
Db    425 RQTSRTTTAVMLYLSLMQPKPGAPRLQPPPNQRGLCSLAADGLWNQKILFEEQDLRKH 484

Qy    595 GLDGAIISTFLKMGILQEH-PIPLSYSFIHLCFQEFFAAMSYVL-EDEKGRGKHSNCIID 652
Db    485 GLDGEDVSAFLNMNIFQKDINCERYSFIHLSFQDFFAAMYYILDBEGEGGAGPDQ----D 540

Qy    653 LEKILBAYGI-HGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ- 708
Db    541 VTRLLTEYAFSERSFLALTSRFLFGLLNEETRSHLEKSLCWKVSPHIKMDLLQWIQSKAQ 600

Qy    709 ---LILQPHSLESLHCLYETRNKTFLTQVMAHFEBMGMC--VETDMELLVCTFCIKPSRHV 764
Db    601 SDGSTLQQSSLRFPSCLIYEIQEBEFIQOALSHFQVIVVSNIASKMEHMVSSPCLKRCRSA 660

Qy    765 KKLQLI-----------EGRQHRSTWSPTMVVLFR--WVPVTDAYWQIIFSVSLVKVTVRNLKELD 814
Db    661 QVLHLYGATYSADGEDRARCSAGAHTLLVQLRPERTVLLDDAYSEHLAAALCTNPNLIELS 720

Qy    815 LSGNSLSHSAVKSLCKTLRRPRCLLETRLAGCCLTAEDCKDLAFGLRANQTLTTELDLSF 874
Db    721 LIYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 780

Qy    875 NVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNL 934
Db    781 NGVGFPGWMLLCEGLRHPQCRLQMIQLRKCQLESGACQEMASVLGTNPHLVELDLTGNAL 840

Qy    935 DDVGVRLLCEGLRHPACKLIRLGL-------DQTTLSDEMRQELRALEQEKPQLLIF 984
Db    841 EDLGLRLLCQGLRHPVCRLRTLWLKICRLITAAACDELASTLSVNQSLREIDLSLNEL--- 897

Qy    985 SRRKPSVMPTEGLDTGEMSNSTSSLKRQRLGSERAASHV------AQANLKLLDVS 1035
Db    898 --GDLGVLLCEGL-------RHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLS 950
```

```
RESULT 3
CIS1_HUMAN

ID   CIS1_HUMAN     STANDARD;      PRT;   1034 AA.
AC   Q96P20; O75434; Q8TCW0; Q8TEU9; Q8WXH9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cold autoinflammatory syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
DE   PYD-containing protein 3) (PYRIN-containing APAF1-like protein 1)
DE   (Angiotensin/vasopressin receptor AII/AVP-like).
GN   CIAS1 OR NALP3 OR PYPAF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS FCAS MET-198; VAL-439
RP   AND GLY-627, AND VARIANT MWS VAL-352.
RX   MEDLINE=21547523; PubMed=11687797;
RA   Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A.,
RA   Kolodner R.D.;
RT   "Mutation of a new gene encoding a putative pyrin-like protein causes
RT   familial cold autoinflammatory syndrome and Muckle-Wells syndrome.";
RL   Nat. Genet. 29:301-305(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=21909508; PubMed=11786556;
RA   Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
RA   Mak S., Lora J.M., Briskin M., Jurman M., Cao J., DiStefano P.S.,
RA   Bertin J.;
RT   "PYPAF1, a PYRIN-containing APAF1-like protein that assembles with ASC
RT   and activates NF-kB.";
RL   J. Biol. Chem. 277:11570-11575(2002).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3), VARIANT MWS MET-198, AND
RP   VARIANTS FCAS/MWS TRP-260 AND PRO-305.
RX   MEDLINE=22241234; PubMed=12355493;
RA   Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
RA   Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
RA   Cotter F.E., Thome M., Hitman G.A., Tschopp J., McDermott M.F.;
RT   "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad
```

```
RT   phenotype including recurrent fever, cold sensitivity, sensorineural
RT   deafness, and AA amyloidosis.";
RL   Arthritis Rheum. 46:2445-2452(2002).
RN   [4]
RP   SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
RC   TISSUE=Blood;
RX   MEDLINE=20499367; PubMed=11042152;
RA   Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA   Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA   Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RT   "Cloning and functional analysis of cDNAs with open reading frames for
RT   300 previously undefined genes expressed in CD34+ hematopoietic
RT   stem/progenitor cells.";
RL   Genome Res. 10:1546-1560(2000).
RN   [5]
RP   VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND
RP   ARG-569, AND VARIANT FCAS/MWS TRP-260.
RX   MEDLINE=21987640; PubMed=11992256;
RA   Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
RA   Vaudour G., Meyrier A., Watts R.A., Scott D.G.I Nicholls A.,
RA   Granel B., Frances C., Garcier F., Edery P., Boulinguez S.,
RA   Domergues J.-P., Delpech M., Grateau G.;
RT   "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
RT   and familial cold urticaria: a novel mutation underlies both
RT   syndromes.";
RA   Am. J. Hum. Genet. 70:1498-1506(2000).
RN   [6]
RP   VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
RP   THR-662, AND TISSUE SPECIFICITY.
RX   MEDLINE=22062556; PubMed=12032915;
RA   Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
RA   Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
RT   "Chronic infantile neurological cutaneous and articular syndrome is
RT   caused by mutations in CIAS1, a gene highly expressed in
RT   polymorphonuclear cells and chondrocytes.";
RA   Am. J. Hum. Genet. 71:198-203(2002).
CC   -!- FUNCTION: May function as a potential inducer of apoptosis.
CC       Interacts selectively with apoptosis-associated specklike protein
CC       containing a CARD domain (ASC). This complex may function as an
CC       upstream activator of NF-kappaB signaling.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=Q96P20-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;
CC       Name=3;
CC         IsoId=Q96P20-3; Sequence=VSP_005519;
CC   -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly
CC       expressed in polymorphonuclear cells, undetectable or expressed
CC       at a lower magnitude in B and T lymphoblasts, respectively. High
CC       level of expression detected in chondrocytes. Low or no expression
CC       in the other tissues tested.
CC   -!- DISEASE: Defects in CIAS1 are a cause of familial cold
CC       autoinflammatory syndrome (FCAS), commonly known as familial cold
CC       urticaria. FCAS is rare autosomal dominant systemic inflammatory
CC       disease characterized by episodes of rash, arthralgia, fever and
CC       conjunctivitis after generalized exposure to cold.
CC   -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome
CC       (MWS), a rare autosomal dominant fever syndrome with episodic
CC       urticaria, arthralgia, amyloidosis and progressive sensorineural
CC       deafness.
CC   -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile
CC       neurologic cutaneous and articular syndrome (CINCA), also known as
CC       neonatal onset multisystem inflammatory disease, or NOMID, a
CC       rare congenital inflammatory disorder characterized by a triad of
CC       neonatal onset of cutaneous symptoms, chronic meningitis, and
CC       joint manifestations with recurrent fever and inflammation.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC   -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts
CC       in positions 893, 918 and 926.
```

```
DR   EMBL; AF410477; AAL33908.1; -.
DR   EMBL; AF427617; AAL33911.1; -.
DR   EMBL; AY051115; AAL12497.1; ALT_INIT.
DR   EMBL; AY051116; AAL12497.1; JOINED.
DR   EMBL; AY051117; AAL12497.1; JOINED.
DR   EMBL; AY056060; AAL12497.1; JOINED.
DR   EMBL; AY056059; AAL12497.1; JOINED.
DR   EMBL; AY051118; AAL12498.1; -.
DR   EMBL; AY051114; AAL12498.1; JOINED.
DR   EMBL; AY051113; AAL12498.1; JOINED.
DR   EMBL; AY051111; AAL12498.1; JOINED.
DR   EMBL; AY051112; AAL12498.1; JOINED.
DR   EMBL; AY051116; AAL12498.1; JOINED.
DR   EMBL; AY051117; AAL12498.1; JOINED.
DR   EMBL; AF420469; AAL65136.1; -.
DR   EMBL; AF468522; AAL78632.1; -.
DR   EMBL; AY092033; AAM14659.1; -.
DR   EMBL; AF418985; AAM14640.2; ALT_INIT.
DR   EMBL; AF054176; AAC39910.1; ALT_FRAME.
DR   Genew; HGNC:16400; CIAS1.
DR   MIM; 606416; -.
DR   MIM; 120100; -.
DR   MIM; 191900; -.
DR   MIM; 607115; -.
DR   GO; GO:0016506; F:apoptosis activator activity; NAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   GO; GO:0006954; P:inflammatory response; IMP.
DR   GO; GO:0007165; P:signal transduction; NAS.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF005660; LRR; 2.
DR   SMART; SM003368; LRR_Ri; 3.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; Repeat; Leucine-rich repeat; Alternative splicing;
KW   Disease mutation; Deafness.
FT   DOMAIN        1     91       DAPIN.
FT   DOMAIN      218    534       NACHT.
FT   REPEAT      738    761       LRR 1.
FT   REPEAT      795    818       LRR 2.
FT   REPEAT      852    875       LRR 3.
FT   REPEAT      881    904       LRR 4.
FT   REPEAT      909    937       LRR 5.
FT   REPEAT      939    961       LRR 6.
FT   REPEAT      966    989       LRR 7.
FT   DOMAIN      688    695       POLY-GLU.
FT   VARSPLIC    718   1034       Missing (in isoform 3).
FT                                /FTId=VSP_005519.
FT   VARSPLIC    719    775       Missing (in isoform 1).
FT                                /FTId=VSP_005520.
FT   VARSPLIC    834    890       Missing (in isoform 1).
FT                                /FTId=VSP_005521.
FT   VARIANT     198    198       V -> M (in FCAS AND MWS).
FT                                /FTId=VAR_013227.
FT   VARIANT     260    260       R -> W (in FCAS AND MWS).
FT                                /FTId=VAR_013104.
FT   VARIANT     303    303       D -> N (in CINCA AND MWS).
FT                                /FTId=VAR_014104.
FT   VARIANT     305    305       L -> P (IN FCAS AND MWS).
FT                                /FTId=VAR_014105.
```

```
FT VARIANT 309 309 /FTId=VAR_014124.
FT                  F -> S (IN CINCA).
FT VARIANT 348 348 /FTId=VAR_014106.
FT                  T -> M (IN MWS).
FT VARIANT 348 348 /FTId=VAR_014366.
FT                  A -> V (IN MWS).
FT VARIANT 352 352 /FTId=VAR_013228.
FT                  H -> R (IN CINCA).
FT VARIANT 358 358 /FTId=VAR_014367.
FT                  T -> N (IN CINCA).
FT VARIANT 436 436 /FTId=VAR_014368.
FT                  A -> T (IN MWS).
FT VARIANT 439 439 /FTId=VAR_014369.
FT                  A -> V (IN FCAS).
FT VARIANT 439 439 /FTId=VAR_013229.
```

```
Query Match   17.0%;  Score 1279;  DB 1;  Length 1034;
Best Local Similarity  33.0%;  Pred. No. 3.2e-74;
Matches 332;  Conservative 155;  Mismatches 353;  Indels 166;  Gaps 17;
```

```
Qy   8  RLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQYGEQRAW   67
Db   7  KLARYLEDLEDVDLLKKFKMFLLEDYPPQKGCIPLPRGQTEKADHVDLATLMIDFNGEEKAW 66

Qy  68  DLALHTWEQMGLRSLCAQAQEGAGHSPSPPYSPSEPHLGSPSQPTSTAVLM---PWIHEL  124
Db  67  AMAVWIFAAINRDLYEKAKR----------------DEPKWGSDNARVSNPTVICQEDSIEEE 114

Qy 125  PAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTSTAVLG  184
Db 115  WMGLLEYLSRISICKMKKDYRKKYRKYVRS-RFQCI------------EDRNARLGESV-- 160

Qy 185  SWGSPPQPSLAPREQEAPGTQWPLDETSGLYYTEIRERERKSEKGRPPWAAVVGTPPQA   244
Db 161  ------------SLNKR------------YTRLRLIKEHRSQQEREQELLAIGKKTYC   194

Qy 245  HTSLQPHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSWPDYV   304
Db 195  ESPVSP----------------------------------------------------  200

Qy 305  EENRGHLIEIRDLFGPGLDTQEP-RIVILQGEAGIGKSTLARQVKEANGRGQLYGDRFQH  363
Db 201  ------IKMELLFDPDDEHSEPVHTVVFQGAAGIGKTILARKMMLDWASGTLYQDRPDY   253

Qy 364  VFFPSCRELAQSKVVSLAEILGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQEPS  423
Db 254  LPYIHCREVSLVIVTQRSLGDLIMSCCPDPNPIHKIVRKPSRILFLMDGFDELQGAPDEHI 313

Qy 424  SELCLHWSQPQPADALIGSLLGKTILPEASFLITARTTALQNTLPSLEQARWEVLGFSE   483
Db 314  GPLCTDWQKAERGDILLSSLIRKKLLPEASLLITTRPVALEKLQHLLDHPRHVEILGFSE   373

Qy 484  SSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKLTLT   543
Db 374  AKRKEYFPRXYFSDEAQARAAFSLIQENEVLFTMCFIPLVCWIVCTGLKQQMESGKSLAQT 433

Qy 544  SKTTTTLCLHYLAQALQAQ-------PLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLD 597
Db 434  SKTTTAVVFFLSSLLQPRGGSQEHGLCAHLWGLLCSLAADGIWNQKILFEESDLRNHGLLQ 493

Qy 598  GAIISTFLKWGILQ-EHPIPLSYSFIHLCPQEFFAAMSVVLEDEK-GRGKHSNCII----  651
Db 494  KADVSAFLRMNLFQKEVDCEKFYSFIHMTFQEFFAAMYLLEEEKGRTNVPGSRLKLPS   553

Qy 652  -DLEKTLEAYG-IHGLFGASTTRFLLGILLSDEGEREMENIFHCRLSQ--GRNLMQWV--  704
Db 554  RDVTVLLENYGKFEKGYLIFVVRFLFGLLVNQERTSYLEKKLSCKISQQIRLELLKWIEVK 613

Qy 705  -PSLQLLLQPHSLESHLCHLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCI----  758
Db 614  AKAKKLQIOPSQLELFYCILYEMQEEDFVQRAMDYFPKIEINLSTRMDHNVSSFCIENCHR 673

Qy 759  ----------------------------KFSRHVKKLQLIEGRQHRSTWSPTHVVLFRWVPVTDAYNQ 798
```

```
Db 674  VESLSLGFLHNMPKEEEEEEEEKEGGRHLDMVQCVLPSSSHAACSHGLVNSH----LTSSFCR 729
Qy 799  ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDLA  858
Db 730  GLFSVLSTSQSLTELDLSDNSLGPDGPMRVLCETLQHPGCNIRLRLWLGRCGLSHECCFDIS 789
Qy 859  FGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVL  918
Db 790  LVLLSSNQKLVELDLSDNALGDFGIRLLCVGLKHLLCNLKKTLWLVSCCLTSACCQDLASVL 849
Qy 919  SASPSLKELDLIQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLS             964
Db 850  STSHSLTRLYVGENALGDSGVAILCEKAKNPQCNLQKLGLVNSGLT             895
```

```
RESULT 4
CIS1_MOUSE
ID   CIS1_MOUSE   STANDARD;   PRT;   1033 AA.
AC   Q8R4B8;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing
DE   APAF1-like protein 1) (Mast cell maturation inducible protein 1).
GN   CIAS1 OR PYPAF1 OR MMIG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/cJ;
RA   Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA   Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA   Chang H.-W., Sawada M., Ingaaki N., Nagai H., Murakami M., Kudo I.;
RT   "Identification of inducible genes during in vitro maturation of mouse
RT   bone marrow-derived mast cells to connective tissue-type mast cells.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: May function as a potential inducer of apoptosis.
    Interacts selectively with apoptosis-associated specklike protein
    containing a CARD domain (ASC). This complex may function as an
    upstream activator of NF-kappaB signaling (By similarity).
-!- SIMILARITY: Contains 1 DAPIN domain.
-!- SIMILARITY: Contains 1 NACHT domain.
-!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
```

```
DR   EMBL; AF486632; AAL90874.1; -.
DR   MGD; MGI:2653833; Cias1.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00368; PAAD_DAPIN; 1.
DR   PROSITE; PS08024; DAPIN; 1.
DR   PROSITE; PS08037; NACHT; 1.
KW   Apoptosis; Repeat; Leucine-rich repeat.
FT   DOMAIN      1     91  DAPIN.
FT   DOMAIN    216    532  NACHT.
FT   REPEAT    737    760  LRR 1.
FT   REPEAT    794    817  LRR 2.
FT   REPEAT    851    874  LRR 3.
FT   REPEAT    880    903  LRR 4.
```

```
FT   REPEAT        908    931       LRR 5.
FT   REPEAT        937    964       LRR 6.
FT   REPEAT        965    988       LRR 7.
SQ   SEQUENCE   1033 AA;  118274 MW;  5924690966B12117 CRC64;

Query Match
Best Local Similarity   16.7%;
Matches  331;  Conservative  164;   Score 1254.5;  DB 1;  Length 1033;
Mismatches  334;   Indels 187;   Gaps  23;

Qy     8  RLACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQYGEQRAW    67
Db     7  KLAQVLEDLEDVDLKFKFMHLLEDYPPEKGCIPVPRGQMEKADHLDLATLMIDFNGEEKAW    66

Qy    68  DLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSPEEPHLGPSQPTSTAVLMPWIHELPAG   127
Db    67  AMAVWIFAAINRRDLWEKAKK--------------DQP------EWNDT               95

Qy   128  CTQGGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTSTAVLGSWG   187
Db    96  CTSHSS-MVCQE--DSLEEEW--------------MGLLG                       118

Qy   188  SPPQPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSEKGRPPWAAVGTPPQAHTS    247
Db   119  YLSRISICKKKK-------DYCKMYRRHVRSRYYSIKDRN---ARLG-               155

Qy   248  LQPHHHPWEPSVRESLCSTWPWKNEDFNQKPTQLILLLQRPHPRSQDPLVKRSWPDYVEEN   307
Db   156  --------------ESVDLNSRYTQLQLV-KEHPSKQE--------                 178

Qy   308  RGHLI-----EIRD-------LFGPGLDTQEP-RIVILQGRAGIGKSTLARQVKEA       350
Db   179  REHELLTIGRTKMRDSPMSSLKLELLFEPEDGHSEPVHTVVFQGAAGIGKTILARKIMLD   238

Qy   351  WGRGQLYGDRPRQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILD   410
Db   239  WALGKLFKDKFDYLFFIHCREVSLRTPRSLADLIVSCWPDDPNPPVCKILRKPSRILFLMD   298

Qy   411  GVDEPGWVLQEPSSELCLHWSQPQPADALLGSSLLGKTILPEASFLITARTTALQNLIPSL   470
Db   299  GFDELQGAFDEHIGEVCTDWQKAVRGDILLSSLIRKLLPKASLLITTRPVALEKLQHLL    358

Qy   471  EQARWVEVLGPSESSRKEYFYRYFDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCL    530
Db   359  DHPRHVBILGFSEAKREYPKYFSNELQAREAFRLIQENEVLFTMCPIPLVCWIVCTGL    418

Qy   531  MQQMRKEKLTLTISKTTTTLCLHYLAQALQAQP-----LGPQLRDLCSLAARGIWQKKT    584
Db   419  KQOMETGKSLAQTSKTTAVFFLSSLLQSRGGIEEHLFSDYLQGLCSLAADGIWNQKI      478

Qy   585  LFSPDDLRKHGLDGAIISTFLKMGILQ-EHPIPLSYSFIHLCPQEFFAAMSYVLEDE---   640
Db   479  LFEEDCLRKGLQKTDVSAFLRMNVPQKEVDCERPYSFSHWTFQBPFAANYLLEEBAG      538

Qy   641  ---KGRGKHSNCI-IDLEKTLEAYG-IHGLFGASTTRFLLGLLSDEGBREMENIFHCRL    694
Db   539  ETVRKGPGCGSDLLNRDVKVLLENYGKFEKGYLIFVVRFLFGIVNQERTSYLEKKLSCKI   598

Qy   695  SQ--GRNLMQWV----PSLQLLLLQPHSLESLHCLYETRNKTFLTQVWAHFEEMGMCVBTD   748
Db   599  SQQVRLELLKWIEVKAKAKLLQWQPSQLELFYCLYEMQEEDFVQSAMDHFPKIEINLSTR    658

Qy   749  MELLVCTFCIKFSRHVKKLQL--------IEGRQHRSTWSPT------MVVLPRWV       790
Db   659  MDHVVSSFCIKNCHRVKTLSLGFFHNSPKEEBEERRGGRPLDQVCVFPDTHVACSSRLV    718

Qy   791  P--VTDAYWQLLFSVLKVTRNLKELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTS   848
Db   719  NCCLTSSFCRGLFSSLSTNRSLTELDLSDNTLGDPGMRVLCEALQHPGCNIQRLWLGRCG    778

Qy   849  LTAEDCKDLAFGLRANQTLTELDLSFNVLTDGAKHLCQRLROPSCKLQRLQLVSCCLTS    908
Db   779  LSHQCCPDISSVLSSQKIVELDLSDNALGDPGIRILLCVGLKHLLCNLQKLWLVSCCLTS   838

Qy   909  DCCQDLASVLSASPSLKEDLLQQNNLDDVGVRLLLCEGLRHPACKLIRLGLDQTTLS      964
Db   839  ACCQDLALVLSSNHSLITRLYIGENALGDSGVQVLCERKKDPQCNLQKIGLVNSGLT      894
```

RESULT 5
NAL4_HUMAN

```
ID   NAL4_HUMAN     STANDARD;     PRT;   994 AA.
AC   Q96MN2; Q9AY6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   NACHT-, LRR- and PYD-containing protein 4 (PAAD and NACHT-containing
DE   protein 2) (PYRIN-containing APAF1-like protein 4) (Ribonuclease
DE   inhibitor 2).
GN   NALP4 OR PYPAF4 OR PAN2 OR RNH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Martinon F., Tschopp J.;
RT   "NALP4 a novel member of the PYD, NACHT, and LRR family.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Fiorentino L., Reed J.C.;
RT   "Pan2, a novel PAAD-containing protein.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Bertin J.;
RT   "PYPAF4: a novel PYRIN-containing APAF1-like protein.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Miyamoto T.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RA   Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA   Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA   Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA   Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 436-994 FROM N.A. (ISOFORM 2).
RC   TISSUE=Placenta;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
```

```
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1;
CC         IsoId=Q96MN2-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q96MN2-2; Sequence=VSP_003916;
CC       Name=3;
CC         IsoId=Q96MN2-3; Sequence=VSP_003917;
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
CC   -!- CAUTION: A stop codon in Ref.4 was read through in position 41 to
CC       extend the sequence.
CC   -----------------------------------------------------------------------
```

```
DR   EMBL; AF442488; AAL35293.1; -.
DR   EMBL; AY072292; AAL68396.1; -.
DR   EMBL; AF479747; AAL87104.1; -.
DR   EMBL; AF482706; AAL88672.1; -.
DR   EMBL; AK056688; BAB71254.1; -.
DR   EMBL; BC016443; AAH16443.1; -.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Repeat; Leucine-rich repeat; Alternative splicing.
FT   DOMAIN        1     94       DAPIN.
FT   DOMAIN      149    472       NACHT.
FT   REPEAT      637    660       LRR 1.
FT   REPEAT      698    721       LRR 2.
FT   REPEAT      722    745       LRR 3.
FT   REPEAT      750    777       LRR 4.
FT   REPEAT      806    833       LRR 5.
FT   REPEAT      863    886       LRR 6.
FT   REPEAT      920    943       LRR 7.
FT   REPEAT      949    972       LRR 8.
FT   VARSPLIC      1     93       Missing (in isoform 2).
FT                                /FTId=VSP_003916.
FT   VARSPLIC    731    786       MAASFFSDFGLMWYLEELKKEEFRKFKEHLKQMTLQELELKQ
FT                                IPWTEVKKASRELANLLIKHYEEQQAWNITLRIFQKMDRK
FT                                DLCMKVMRERT -> MQBCLTLWVFSPLALTDS (in
FT                                isoform 3).
FT                                /FTId=VSP_003917.
FT   CONFLICT     55     55       L -> P (IN REF. 4).
FT   CONFLICT     99     99       Y -> H (IN REF. 4).
FT   CONFLICT    177    177       I -> T (IN REF. 4).
FT   CONFLICT    209    209       W -> R (IN REF. 4).
FT   CONFLICT    296    296       I -> V (IN REF. 4).
FT   CONFLICT    720    720       Y -> C (IN REF. 4).
FT   CONFLICT    957    957       Y -> D (IN REF. 4).
SQ   SEQUENCE   994 AA;  113444 MW;  1CECFC68B1B398C9 CRC64;
```

```
Query Match    13.8%;  Score 1037;  DB 1;  Length 994;
Best Local Similarity  25.5%;  Pred. No. 1.1e-58;
Matches  303;  Conservative 168;  Mismatches 313;  Indels 402;  Gaps 25;

Qy     9  LACYLEFLKKEELKEFQLLANKAHSRSSS-----GETP-AQPEKTSGMEVASYLVAQYG  62
          |||||| ||||||| :||         ::|   | ::|
Db    11  LMWYLEELKKEEFRKF------KEHLKQMTLQELELKQIPWTEVKKASRELANLLIKHYE  64

Qy    63  EQRANDLALHTWBQMGLRSLCAQA-QEGAGHSPSFPYSPSBPHLGSSPQPTSTAVLMPWI  121
          | |:|                        ::             :
Db    65  BQQAWNITLRIFQKMDRKDLCMKVMRERTGYTKTY----  99

Qy   122  HELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHBSPSQBSSPNAPTSTA  181
                           |||||
Db   100  ------QAHAQKFSRL-  110

Qy   182  VIGSWGSPPQPSLAPREQBAPGTQWPLDETSGIYYTEIRERBEREKSGRPPWAAVVGTP  241
                                                 |||
Db   111  ----NSSKS-  119

Qy   242  PQAHTSLQPHHHPWEFPSVRBSLCSTWPWKDNEDFNQKFTQLILLLQRPHPRSQDPLVKRSWP  301
                               : |:                     ::|
Db   120  -HLYFEEEVKQBEC-  132

Qy   302  DYVEENRGHLLEIRDLFGPGLDTQEFPRIVILQGAAGIGKSTLARQVKEAWRGQGLYGDRF  361
           |::|                    ::      |||    | :
Db   133  DHLDR------LFAPKETGKQPRTVIIQGPGGIGKTTLLMKLMMAWSDNKIFRDRF  182

Qy   362  QHVFYFSCRELAQSKVVSILAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQE  421
          |:|| ||||| :| | :| |        |  |:||   |||||||  | ||||
Db   183  LYTFYPCCRELRELPPTSLADLISRRWPDPAAPITEIVSQPERLLFVIDSFEBELQGGLNE  242

Qy   422  PSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWEVLGF  481
           ::| | ::  || :||| ||| ||  |||| ||| | :  ||:||||
Db   243  PDSDLCGDLMEKRPVQVLLSSLLRKKMLPEASLLIAIKPVCPKELRDQVTISEIYQPRGF  302

Qy   482  SESSRKEYFRYFTDERQAIRAFRLVKSNKELWALCLVPWSWLACTCLMQQMKRKEKLT  541
           :|::|:|| :|   : | || ||| ::| || | :|      |  |:||
Db   303  NESDRLVYFCCFFKDPKRAMEAFNLVRBESEQLFSICQIPLLCWILCTSLKQEMQKGKDLA  362

Qy   542  LTSKTTTTLCLHYLAAQALQQPL-GP-----QRDLCSLAAEGIWQKKTLFSPDDLRKH  594
          || :||:|   | :::|| | |    |     ||||||||| |   |||||||| ||
Db   363  LTCQSTTSTVVSSFVFNLPTPEGAEGPTPQOTQHOLKALCSLAAABGWMWTDTFBFCEDDLRRN  422

Qy   595  GLDGALISTFLKMGILQEH-PIPLSYSFIHLCFPQEFFAAMSYVLEDEKGRGKHSN---C  649
          | :|  | |  |  ||:|  || ||| ||: | |||| ||:|:|:|||| ||:|   |
Db   423  GVVDADIPALLGTKILLKYGERBSSYVFLHVCIQEFCAALFYLL---KSHLDHPHAVRC  479

Qy   650  IIDL-----EKTLEAYGHGLFGASTTRFLLGLLSDEGERBMENIFHCRLSQ--GRNLMQ  702
          |  |      | ||    ::  |     |:|| ||| : | | :|  :|:|   ||::|
Db   480  VQELLVANFEKARRAHWI--FLGC----FLTGLLNKKEQKLDAFFGFQLSQEBIKQQIHQ  533

Qy   703  WVPSLQLLLQPH------SLBSLHCLYETRNKTFLTQVMAHFBEMGMCVETDMELLVCTFCI  758
           : || ||| |       ||:||||| :|  |||| |:| ||:|||| |::|::|| | |
Db   534  CLKSLGERGNPQGQVDSLAIFYCLFEMQDPAFVKQAVNLLQBANFHIIDNVDLLVSAYCL  593

Qy   759  KF-----SRHVKKLQLIEGRQHRS  777
          |:      :|::|:| |:|||:|
Db   594  KYCSSLRKLCFSVQNVFKBEDEHSSTSDYSLICWHHICSVLTTSGHLRELQVDSTLSES  653

Qy   778  T------WSPTMVVLFRWVPVTIDAYWQILFSVLKVTR----  808
                     ||  ::| |    |  |:| |:| ||| ||:
Db   654  TFVTWCNQLRHPSCRLQKLGINNVSFSGQSVLLFEVLFYQPDLKYLSFTLFKLSRDIRS  713

Qy   809  IIDL------NLKE--  812
          | |        ||
Db   714  LCDALNYPAGNVKELALVNCHLSPIDCEVLAGLLTNNKKLTYLNVSCNQLDTGVPLLCEA  773

Qy   813  -----LDLSGNSLSHSAVKSLCKTLRR  834
               ||||||||:|:||| |||||||
Db   774  LCSPDTVLVYLMLAFCHLSEQCCEYISEMLLRNKSVRYLDLSANVLKDBGLKTLCEALKH  833

Qy   835  PRCLLBTIRILAGCGLTABDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSC  894
           ||  |  :|: ||| |: ||:||: | | ||||  | || ::|: |   || |||:: |
Db   834  PDCCLDSLCLVKCFITPAAGCBDLASALISNQNLKILQIGCNEIGDVGVQOLLCRALTHTDC  893

Qy   895  KLQRLQLVSGCLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLLCBGLRHPACKLI  954
           ||| | |  ||||:|| |||||| :|| ||::||  |||| ||||:|| | |||| | |
Db   894  RLEILGLEBCGLTSTCCKDLASVLTCSKTLQQLNLTLNTLDHTGVVVLCEALRHPBCALQ  953

Qy   955  RLGLDQTTLSDBMRQELRALEQEKPQLLIFSRRKPSVMTPFEGLDT  1000
```

```
Db

RESULT 6

Db   MATE_HUMAN                                      954 VLGLRKTDFDEETQALITAEEERNPNLTI-------TDDCDT 988
                                                         |||||||||||||||||||||||||||||       |||||||

ID   MATE_HUMAN              STANDARD;      PRT;   1200 AA.
AC   P59047;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mater protein homolog.
GN   MATER.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=21922687; PubMed=11925379;
RA   Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT   "A human homologue of mouse Mater, a maternal effect gene essential
RT   for early embryonic development.";
RL   Hum. Reprod. 17:903-911(2002).
CC   -!- TISSUE SPECIFICITY: Oocyte-specific.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC   -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC       Mater.
```

```
DR   EMBL; AY054986; AAL15549.1; -.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00560; LRR; 3.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00368; LRR_RI; 11.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN       57    148       DAPIN.
FT   DOMAIN      280    602       NACHT.
FT   REPEAT      704    727       LRR 1.
FT   REPEAT      730    753       LRR 2.
FT   REPEAT      780    803       LRR 3.
FT   REPEAT      809    832       LRR 4.
FT   REPEAT      836    863       LRR 5.
FT   REPEAT      865    892       LRR 6.
FT   REPEAT      893    916       LRR 7.
FT   REPEAT      950    973       LRR 8.
FT   REPEAT      979   1002       LRR 9.
FT   REPEAT     1007   1034       LRR 10.
FT   REPEAT     1036   1059       LRR 11.
FT   REPEAT     1064   1092       LRR 12.
FT   REPEAT     1121   1142       LRR 13.
FT   NP_BIND     286    293       ATP (POTENTIAL).
SQ   SEQUENCE   1200 AA;  134235 MW;  9A07D02A771B28FA CRC64;
```

Matches 329; Conservative 176; Mismatches 411; Indels 380; Gaps 39;

```
Qy     5 AWGRLACYLPFLKKEELKFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ-YGE  63
Db    62 SYGLQWCLYE-LDKEEFQTIKELLKKKS-SESTTCSIPQFEIENANVECLALLLHEYGA 119

Qy    64 QRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPWIHE 123
Db   120 SLAWATSISIPENMNLRTLSEKARDDMKR-----HSPEDPEATMTDQGPS-----KE   166

Qy   124 LPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSPSQESPNAPTSTAVL 183
Db   167 KVPGISQAVQQ-----------------------DSATAAETKEQEISQA-           193

Qy   184 GSWGSPPQPSLAPRRQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGTPPQ  243
Db   194 -------MEQBGA----------------------TAABTEEQEISQAMEBQGATAAETEEQ 226

Qy   244 AHTSLQPHHIPWEPSVRESLCSTWPWKNEDNQKFTQLLILQRPHPRSQDDPLVKRS---  299
Db   227 GHGG------------------------------DTWDYKSHVWT-KF-            256

Qy   300 --WPDYVEENRGHLIEIRDLFGPGLDTQ-----EPRIVILQGAAGIGKSTILARQVKEAWG 352
Db   257 AADWP--------EMQTLAG-AFDSDRWGFRPRTVVLHGKSGIGKSALARRIVLCWA     304

Qy   353 RGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATAPAPIRQILSRPERLLFILDGV 412
Db   305 QGGLYQGMPSYVFFLPVRRMQRKKESSVIEFISREWPDSQAPVTEIMSRPERLLFTIDGF  364

Qy   413 DEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKITILPEASFLITARTTALQNLIPSLEQ 472
Db   365 DDLGSVLNN-DTKLCKDWAEKQPPFTLIRSLLRKVLLPESFLIVTVRDVGTEKLKSEVVS  423

Qy   473 ARWEEVLIGFSESSRKEYFYRYFTDERQAIRAPRLVKSNKELWALCLVPWYSWLACTCLMQ 532
Db   424 PRYLLVRGISGEQRIHLLLERGIGEHQKTQGLRAIINNRELLDDQCCVPAVSSLICVALQL  483

Qy   533 QMKRKEXLTLITSKTTTTLCLHYLAQALQAQPLG----------POLRDLCSLAAEGIWQ  581
Db   484 QDVVGESVAFNQTLTG--LHAAFAFHQLTPRGVVRRCLNLEERVVLKRFCRMAVEGVWN   541

Qy   542 RKSVFPDGDLMVQGLGGESELRALFHMNIL---LFDSHCEEYYTFFHLSLQDFCAALYYV  597
Db   598 LEGLE--IEPALCPLYVEKTKRSMELKQAGFHIHSLW---MKRFLPGLVSEDVRRPLEVL  652

Qy   582 KKTLFSPDDLRKHGLDGAIISTFLRKNGILQEHPIPLS-----YSFFHLCFQEFFAMSYV  636
Db   365 ...

Qy   637 LEDEKGRGKHSNCIIDLEKTLEA-------YGIHGLFGASTTRFLLGLLSDEGEREMENI  689
Db   653 LGCPVPLGVRQKLLHWVSLLG---QQPNATTPGDTLDAFHCLPETQDKEFVRLALNSFQEV 710

Qy   690 FHCRLSQG--RNLMQWVPSLQLLLQPH----------SLRSLHCLYETRNKTFLTQVMAHPEEM 741
Db   711 WLPINQNLDIIASSFCLQHCPYLRKIRVDVAGIPPRDESAEACPVVPLWMRDKTLIEEQW   770

Qy   742 GMCXVETDMELLVCTFCIKFSRHVKKLQL-IEG----RQHRSTWSPTMVVLFRWVPVTDAYW 797
Db   771 EDFCSMLGTIHPHLRQLDIGSSILTERAMKTLCAKLRHPTCKIQTLMFRNAQITPGVQHLW  830

Qy   798 ----QILF  801
Db   802 SVLKYTRNLKELDLSGNSLSHSAAVKSLCKTIRRPRCLLETLRLAGGGLT----------  850

Qy   831 RIVMANRNLRSLNLGGTHLKBEDVRMACEBALKHPKCLLESLRLDCCGLTHACYLKISQIL 890
Db   851 ---------------------------------------------------AED-----   853

Qy   851 ---------------------------------------------------AED-----  853
Db   854 ---CKDLAFGLRANQTL  867

Qy   891 TTSPSLKSLSLAGNKVTDQGVTPLSDALRVSQCALQKLILEDCGITATGCQSLASALVSN  950
Db   891 ...

Qy   951 RSLITHCLSNNSLGNEGVNLLCRSMRLPHCSLQRLMLNQCHLDTACGSLALALMGNSWL 1010
Db   951 ...
```

Best Local Similarity 13.0%;

Query Match 25.4%;

Score 983; DB 1; Length 1200;

Pred. No. 4.3e-55;

```
Qy    868  TELDLSFNVLIDAGAKHLCQRLRQPSCKLQRLQVSCGLITSDCCQDLIASVLSASPSLKEL  927
Db   1011  THLSLSMNPVEDNGVVKLLCEVMREPSCHLQDLELVKCHLITAACCESLSCVISRSRHLKSL 1070

Qy    928  DLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALRQEGKPQLLIFSRR  987
Db   1071  DLTDNALGDGGVAALCEGLKQKNSVLTRLGLKACGLITSDCCEALSL-------ALSCNRH 1123

Qy    988  KPSVMTPIEGLDTGEMSNSTSSLKRQLGSERAASHVAQANLKLLDVSK-IFP--IAEIA 1044
Db   1124  LTSU--------------NLVQNNFSPKGMWKLCSAPAC---PTSNLQIIGLWKWQYPVQIRKLL 1171

Qy   1045  EESSPEVVPVELLCVPSPASQGDLHTKPLGTGTDDDPW 1080
Db   1172  EE-------VQLL-KPRVVIDGSWHS--FDEDDRYW 1197
```

```
RESULT 7
ID   NAL2_HUMAN      STANDARD;      PRT;   1062 AA.
AC   Q9NX02; Q9BVN5; Q9HAV9; Q9NWK3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   NACHT-, LRR- and PYD-containing protein 2 (Nucleotide-binding site
DE   protein 1) (PYRIN-containing APAF1-like protein 2).
GN   NALP2 OR NBS1 OR PYPAF2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21169419; PubMed=11027363;
RA   Bertin J., DiStefano P.S.;
RT   "The PYRIN domain: a novel motif found in apoptosis and inflammation
RT   proteins.";
RL   Cell Death Differ. 7:1273-1274(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1)
RX   MEDLINE=21148093; PubMed=11250163;
RA   Martinon F., Hofmann K., Tschopp J.;
RT   "The pyrin domain: a possible member of the death domain-fold family
RT   implicated in apoptosis and inflammation.";
RL   Curr. Biol. 11:R118-R120(2001).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2)
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Colon, and Kidney epithelium;
RA   Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA   Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA   Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Lung, and Placenta;
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be implicated in apoptosis (By similarity).
CC   -!- COFACTOR: Binds ATP (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9NX02-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9NX02-2; Sequence=VSP_005522;
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
```

```
DR   EMBL; AF298547; AAG15253.1; ALT_INIT.
DR   EMBL; AF310106; AAG30289.1; -.
DR   EMBL; AF464764; AAL69962.1; -.
DR   EMBL; AK000517; BAA91223.1; -.
DR   EMBL; AK025952; BAB15293.1; -.
DR   EMBL; AK007084; BAA91377.1; ALT_INIT.
DR   EMBL; BC003592; AAH03592.1; -.
DR   EMBL; BC001039; AAH01039.1; -.
DR   HSSP; P10775; 2BNH.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW   Alternative splicing.
FT   DOMAIN        1     94       DAPIN.
FT   DOMAIN      207    526       NACHT.
FT   REPEAT      467    491       LRR 1.
FT   REPEAT      622    645       LRR 2.
FT   REPEAT      754    777       LRR 3.
FT   REPEAT      810    832       LRR 4.
FT   REPEAT      839    862       LRR 5.
FT   REPEAT      867    890       LRR 6.
FT   REPEAT      924    947       LRR 7.
FT   REPEAT      981   1005       LRR 8.
FT   REPEAT     1010   1033       LRR 9.
FT   NP_BIND     213    220       ATP (POTENTIAL).
FT   DOMAIN      518    523       POLY-GLU.
FT   VARSPLIC    133    154       Missing (in isoform 2).
FT                                /FTId=VSP_005522.
FT   CONFLICT      1      1       M -> V (IN REF. 1).
FT   CONFLICT     35     35       M -> P (IN REF. 1).
FT   CONFLICT    304    304       I -> S (IN REF. 4;
FT                                BAB15293).
FT   CONFLICT    364    364       R -> K (IN REF. 1).
FT   CONFLICT    980    980       MISSING (IN REF. 1).
FT   CONFLICT   1052   1052       A -> E (IN REF. 1).
SQ   SEQUENCE   1062 AA;  120514 MW;  4DBB0F6E9CCBC8A7 CRC64;
```

```
Query Match    12.6%; Score 948; DB 1; Length 1062;
Best Local Similarity 27.8%; Pred. No. 6.4e-53;
Matches 302; Conservative 167; Mismatches 389; Indels 228; Gaps 32;

Qy   148 WREISASLLYQALPSSPDHESPSQESPNAPTSTAVLGSWGSPPQPSLAPREQEAPGTQWP 207
Db    67 WVEMASLQVPEKM-----HRMDLSERADEVREAALKSFNKRKPLSLGITTRKERP----P 117

Qy   208 LDETSGIYYTEIREREREKSEKGRPPWAAVVGTPPQAHTSLQPH-----HHPWEPSVRES 262
Db   118 LD------VDEMLERFKTEAQAFTETKGNVICLGKEVFGGKKPDKDNRCRYILKTKFRE- 170

Qy   263 LCSTWPWKNEDFN---QKPTQLILLQRPHPRSQODPLVKRSWPDYVEENRGHLIEIRDLFG 319
Db   171 MWKSWPGDSKEVQVMAERYKMLIPFSNP------------------------RVLPG 203

Qy   320 PGLDTQEPRIVILQGSAAGIGKSTLARQVKEAWGRGQLYGDRFQRVYFSCRELAQSKVVS 379
Db   204 PFSYT-----VVLYGPALGLGKKTLLAQKLMLDWAEDNLI-HKFKYAFYLSCRELSRLGPCS 257

Qy   380 LAELLGKDGTATPAPIRQILSRPERLLFILDGVDE-----PGWVLQEPSSELCLHWSQPQP 435
Db   258 FAELVFRDWPELQQDDIPHILAQARKILFVIDGFDELGAAPGALIE----DICGDWEKKKP 313

Qy   436 ADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFT 495
Db   314 VPVLLGSLLNRVMLPKAALLVTTRPRALRDLRILAEEPIYIRVEGFLEEDRRAYFLRHFG 373

Qy   496 DERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKREKLLTLSKTTTTLCLHYL 555
Db   374 DEDQAMRAFELMRSNAALPQLGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFL 433

Qy   556 AQAL-QAQPLGPQLRDLCSLLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHP 614
Db   434 CSRRPQGAQLRGALRTLSLLAAQGIWAQTSVLHREDLERLGVQESDLRLFLDGDILRQDR 493

Qy   615 IPLS-YSFIHLCTQEFPRAAMSYVL--EDEKGRGKHSNCIIDLEKTLEAYGIHGLFGASTT 671
Db   494 VSKGCYSFIHLSFQQFTALFYTLEKEEEEDRDGHTWDIGDVQKLLS--GVERLRNPDLI 551

Qy   672 R---FLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSL-QLLLQ------PHSL- 716
Db   552 QAGYYSFGLANEKRAKELEATFGCRMS-----------PDIKQELLRCDISCKGHSTVTDL 602

Qy   717 -ESLHCLYEYRNKTFLTQVMAHFSEMGMCVETDMELLVCTFCIKFSRHVKKLQL------ 769
Db   603 QELLGCLYESQEELVKEWMAQFKEISLHLNA-VDVVPSSFCVKHCRNLQKMSLQVIKEN 661

Qy   770 ------------IEGRQ--------------------------HRSTWSPTMV-----784
Db   662 LPENVTASESDAEVERSQDDQHMLFPWTDLCSIFGSNKDLMGLAINDSFLSASLVRILCE 721

Qy   785 VLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLR 833
Db   722 QIASDTCHLQRVFKNISPADAHRNLCLA-LRGHKTVTYLTLQGND-QDDMFPALCEVLR 779

Qy   834 RPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCORLRQPS 893
Db   780 HPECNLRYLGLVSCSATTQQWADLSLALEVNQSLTCVNLSDNELLDEGAKLLYTTLRHPK 839

Qy   894 CKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKL 953
Db   840 CFLQRLSLENCLTEANCKDLAAVLVVSRELTHLCLAKPIGNTGVKFLCEGLRYPECKL 899

Qy   954 IRLGLDQTTLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPEGLDTGEMSNSTSSLKRQ 1013
Db   900 QTLVLWNCDITSDGCCDLTKLLQEKSSLLC------------------------LDLG 933

Qy  1014 RLGSERAASHVAQANLKLLDVSKIFPIAEI---AEESSPEVVPVELLCVPSPASQG--- 1066
Db   934 ------LNHIGVKGMKFLCEALRKPLCNLRCLWLMGCSIPPFSCEDLCGSALSCNQSLVT 986

Qy  1067 -DLHTKPLGTD----------------------DDFWGPTGPVATEVVDKEKNLY---R 1099
Db   987 LDLGQNPLGSSGVRMLFETLTCSSGTLRTLRLKIDDFNDELNKILEEIEEKNPQLIIDTE 1046

Qy  1100 VHFPVA 1105
Db  1047 KHHPWA 1052
```

RESULT 8

PYA3_HUMAN

```
ID   PYA3_HUMAN     STANDARD;      PRT;   980 AA.
AC   Q8WX94;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   PYRIN-containing APAF1-like protein 3.
GN   PYPAF3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF464765; AAL69963.1; -.
DR   InterPro; IPR007091; LRR_RInih.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN        1     93       DAPIN.
FT   DOMAIN      172    491       NACHT.
FT   REPEAT      614    638       LRR 1.
FT   REPEAT      674    697       LRR 2.
FT   REPEAT      760    784       LRR 3.
FT   REPEAT      788    810       LRR 4.
FT   REPEAT      817    840       LRR 5.
FT   REPEAT      845    868       LRR 6.
FT   REPEAT      874    897       LRR 7.
FT   REPEAT      902    928       LRR 8.
FT   REPEAT      933    957       LRR 9.
FT   NP_BIND     178    185       ATP (POTENTIAL).
SQ   SEQUENCE   980 AA;  111806 MW;  822AF2FD43380D3D CRC64;

Query Match    12.1%; Score 908.5; DB 1; Length 980;
Best Local Similarity 30.2%; Pred. No. 2e-50;
Matches 261; Conservative 128; Mismatches 284; Indels 191; Gaps 21;

Qy   294 PLVKRSWPDYVEEN-------RGHLIEIRD-----LFGPGLDTQEPR----IVIL 332
Db   117 PGEKEGWRNSMEKQSLVWKNTFWQGDIDNFHDDVTLRNQFFIPFLNPRTPRKLTPYTVVL 176
```

```
RESULT 9
PYA5_HUMAN

ID    PYA5_HUMAN    STANDARD;    PRT;    892 AA.
AC    P59044;
DT    28-FEB-2003 (Rel. 41, Created)
DT    28-FEB-2003 (Rel. 41, Last sequence update)
DT    15-SEP-2003 (Rel. 42, Last annotation update)
DE    PYRIN-containing APAF1-like protein 5.
GN    PYPAF5.
OS    Homo sapiens (Human).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX    NCBI_TaxID=9606;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=22162427; PubMed=12019269;
RA    Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA    Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT    "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT    activation of NF-kappa B and caspase-1-dependent cytokine
RT    processing.";
RL    J. Biol. Chem. 277;29874-29880(2002).
RN    [2]
RP    FUNCTION.
RX    MEDLINE=22275822; PubMed=12387869;
RA    Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA    Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA    Distefano P.S., Bertin J.;
RT    "Functional screening of five PYPAF family members identifies PYPAF5
RT    as a novel regulator of NF-kappaB and caspase-1.";
RL    FEBS Lett. 530:73-78(2002).
CC    -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC        activation of NF-kappa-B.
CC    -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC    -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC    -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
CC        much lower levels in T-cells.
CC    -!- SIMILARITY: Contains 1 DAPIN domain.
CC    -!- SIMILARITY: Contains 1 NACHT domain.
CC    -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC    --------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    --------------------------------------------------------------------------
DR    EMBL; AF479748; AAL87105.1; -.
DR    InterPro; IPR007091; LRR_RInh.
DR    InterPro; IPR003590; LRR_RInh_sub.
DR    InterPro; IPR007111; NACHT_NTPase.
DR    InterPro; IPR004020; PAAD_DAPIN; 1.
DR    Pfam; PF02758; PAAD_DAPIN_dom.
DR    SMART; SM00368; LRR_RI; 3.
DR    PROSITE; PS50824; DAPIN; 1.
DR    PROSITE; PS50837; NACHT; 1.
KW    ATP-binding; Leucine-rich repeat; Repeat.
FT    DOMAIN      1    103    DAPIN.
FT    DOMAIN    196    513    NACHT.
FT    DOMAIN    604    614    POLY-GLU.
FT    REPEAT    462    487    LRR 1.
FT    REPEAT    727    747    LRR 2.
FT    REPEAT    755    778    LRR 3.
FT    REPEAT    811    834    LRR 4.
FT    REPEAT    845    868    LRR 5.
FT    NP_BIND   202    209    ATP (POTENTIAL).
SQ    SEQUENCE   892 AA;   98733 MW;   4AA8D1FC766DDE9D CRC64;

Query Match            12.0%;   Score 901.5;   DB 1;   Length 892;
Best Local Similarity  27.8%;   Pred. No. 4.8e-50;
Matches    295;   Conservative   143;   Mismatches 328;   Indels 295;   Gaps 33;

Qy      9  LACYLEFLKKEELKEFQLLLANKAHSRSSSGET-P-AQPEKTSGMEVASYLVAQYGEQRA   66
           :|::|:.|.|::||::|| ::|           | | :|      |::       |::
Db     22  LLAALEELSQEQLKRFR---HKLRDVGPDGRSIPWGRLERADAVDLIAEQLAQFYGPEPA   77

Qy     67  WDLALHTWEQMGLRSLCAQAQE-----GAGHSFSPPYSPSEPHLGSPSQPTSTAVLMPW   120
           ::|| : .|||:| :                :||                  :
Db     78  LEVARKTLKRADARDVAAQLQERRLQRLGLSGGSTLLSVSEYK----------------  119

Qy    121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSPSQBSSPNAPTST  180
           :|:| |     ||::||:|
Db    120  --------KKYREHVLQLHARVEBRNARSVKITRRFTKLLIAPE----------------  155

Qy    181  AVLGSWGSPGPQPSILAPREQEAPGTQWPLDETSGIYYTEIRBEREREKSEKGRPPWAAVVGT  240
                      :.|.||                                  ::
Db    156  ------SAAPEEALGPAEPEPG--------------------RARRSD----------  178

Qy    333  QGAAGIGKSTLIARQVKEANGRGQLYGDRPQHVFYFSRCRELAQSKVVSLAELIGKDGTAPP  392
           :|.||:||:|:|: ::||:
Db    177  HGPAGVKTTLAKKCMLDWTDCNI-SPTLRYAFYLSCKELSRMGPCSFABLISKDWPELQ    235

Qy    393  APIRQILSRPERLLFILDGVDE------PGWVLQEPSSELCLHWSQPQPADALLGSLLGKTI  448
           . |::||:: :||||::||:||       ..::|           |||::|||||| :.::
Db    236  DDIPSILAQAQRILFVVDGLDELKVPPGALIQ----DICGDWEKKKPVPVLLGSLLRRKM   291

Qy    449  LPEASFLITARTFALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRLVK  508
           || |::::|:|::||::| : : .:  | || |: .||:|:| ||:| ||::|||||::
Db    292  LPPAALIVTTRPRALRDLQLLAQQPIYVRVEGFLEBDRAYFLRHFGDEDQAWRAFELMR   351

Qy    509  SNKELWALCLIVPWVSWLACTCLMQQMKRKEKLTLTSKTTTLLCHYLLAQAL-QAQPLGPQ  567
           || |: .| .| |  |::.| | |:: ::: .::    |:  : . |:::|  |||:.
Db    352  SNAALFQLGSAPAVCWIVCTTLKLQMEKGEDPVPTCLTRTGLFLRFLCSRPPQGAQLRGA   411

Qy    568  LRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPLS-YSFIHLCF  626
           || | .||| |:| :    ::::::  |: . ::  ::  ::|:: : .    |||||.|
Db    412  LRTLSLLAAQGLWAQMSVPHREDLERLGVQOESDLRLFLDGDILRQDRVSKGCYSFIHLSF  471

Qy    627  QBFPAAMSYVLEDEKG--RGKHSNCIIDLEKTLEAYGIHGLFGASTTR---FLLGLLSDE  681
           | |::|: ::.|::::   |.:.  : : :. :||  : .:   : :     ||.||:.|
Db    472  QQFLTALFYALEKEEGEDRDGHAWDIGDVQKLLS--GEERLKNPDLIQVGHFLFGLANEK  529

Qy    682  GEREMENIFHCRLSQGRNLMQWPSL-QLLLQ------PHSL----ESLHCLYETRN  727
                : :| | ::   | ::|::   ::::       ::::     ||||::| .:
Db    530  RAKELEATFGCRMS-------PDIKQELLQCKAHLHANKPLSVTDLKEVLGCLYESQE  580

Qy    728  KTFLTQVWAHFEEMGMCVETDMELLLVCTPCIKFSRHVKKLQL------------  769
           : ::. : | |::: :.   .: :  | :  ::                ::::
Db    581  EELAKVVVAPFKEISIHLTNTSEVWHCSFSLKHCQDLQKLSLQVAKGVFLENYMDFELDI  640

Qy    770  ------------------------------IEGRQH----------  775
                                         :|||:|
Db    641  EFERCTYLTIPNWARQDLRSLRLWTDFCSLFSSNSNLKFLEVKQSFLSDSSVRILCDHVT  700

Qy    776  RST---------------------------WSPTMVVL------  786
           |||                           :|.:.:
Db    701  RSTCHLQKVEIKNVTPDTAYRDFCLAFIGKKTLTHLTLAGHIEWERTWMLMLCDLLRNHK  760

Qy    787  --FRWVPV-----------TDAYWQILFSVLKVTRNLKELDLSGNSLSHSSAVKSLCKTLRRPCLL  839
             |.|.|            | || ||:| |||.| ::::||||:||:|.|:|:|||| ||||| ||
Db    761  CNL-QYLRLGGHCATPEQWABFFYLKANQSLKHRLSANVLLDEGAMLLYKTMTRPKHFL  820

Qy    840  ETLRLAGCGLTABDCKDLAFGLRANQTLTELDLSPNVLTDAGAKHLCQRLRQPSCKLQRL  899
           :.|.|..| || ::|||||: | ||::|| ||||||:|| ||.|.||| |.:|:|||  |
Db    821  QMLSLENCRLTEASCKDLAAVLVVSKKLTHLCLAKNPIGDTGVKFLCEGLSYPDCKLQTL  880

Qy    900  QLVSCGLTSDCCQDLASVLSASPSLKEILLQQNNLDDVGVRLLCEGLRHPACKLIRLGLD  959
           .|. | ||. .|:||::.|  : |||:|  ::::|:::|:::||: ||:.| |||:::
Db    881  VLQQCSITKLGCRYLSEALQBACSLTNDLSINQIAR-GLWILCQQALENPNCNLKHLRIK  939

Qy    960  QTTLSDEMRQELRALEQEKPQLLI  983
           .|.|::|:|:|::::|::::|:| |
Db    940  TYETNLEIKLLEEVKEKNPKLTI  963
```

```
RESULT 10
PYA5_RAT
ID   PYA5_RAT     STANDARD;     PRT;   854 AA.
AC   Q63035;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYRIN-containing APAF1-like protein 5-like (Angiotensin II/vasopressin
DE   receptor).
GN   PYPAF5 OR AVR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
```

RP   CONCEPTUAL TRANSLATION OF 1-343.
RA   Hinz U.;
RL   Unpublished observations (FEB-2003).
RN   [2]
RP   SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
RC   TISSUE=Kidney;
RX   MEDLINE=96071640; PubMed=7489366;
RA   Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
RT   "Identification of a novel dual angiotensin II/vasopressin receptor on
RT   the basis of molecular recognition theory.";
RL   Nat. Med. 1:1074-1081(1995).
RN   [3]
RP   VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
RX   MEDLINE=21980185; PubMed=11984003;
RA   Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT   "The dual AngII/AVP receptor gene N119S/C163R variant exhibits
RT   sodium-induced dysfunction and cosegregates with salt-sensitive
RT   hypertension in the Dahl salt-sensitive hypertensive rat model.";
RL   Mol. Med. 8:24-32(2002).
RN   [4]
RP   IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA   Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT   "Identification of mammalian orthologs associates PYPAF5 with distinct
RT   functional roles.";
RL   FEBS Lett. 538:173-177(2003).
CC   -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC   activation of NF-kappa-B (By similarity). Angiotensin II and
CC   vasopressin binding protein. May stimulate cAMP accumulation.
CC   -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
CC   levels in all tissues tested.
CC   -!- DISEASE: Defects in PYPAF5 may be a cause of salt-sensitive
CC   hypertension.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC   -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC   sequences, in analogy to ortholog sequences.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; M85183; AAA03623.1; ALT_INIT.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   Pfam; PF00560; LRR; 2.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN   1   103   DAPIN.
FT   DOMAIN   168   484   NACHT.
FT   DOMAIN   564   604   ASP/GLU-RICH.
FT   DOMAIN   655   662   POLY-LYS.
FT   REPEAT   433   458   LRR 1.
FT   REPEAT   610   633   LRR 2.
FT   REPEAT   722   745   LRR 3.
FT   REPEAT   812   836   LRR 4.
FT   NP_BIND   174   181   ATP (POTENTIAL).
FT   VARIANT   492   492   N -> S (IN SALT-SENSITIVE HYPERTENSION).
FT   VARIANT   536   536   C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT   MUTAGEN   766   766   E->K; ABOLISHES ANGIOTENSIN II
FT   BINDING.
SQ   SEQUENCE   854 AA; 95292 MW; D7BBD922D77B3734 CRC64;

Query Match   11.9%;   Score 895; DB 1; Length 854;
Best Local Similarity   33.2%;   Pred. No. 1.2e-49;
Matches 261; Conservative 118; Mismatches 298; Indels 110; Gaps 22;

```
Qy    241  PPQAHTSLQPHHHPWEPSVRESLLCSTWFWKNEDFPNQKFTQLLLLQRPHPRSQDPLVKRSW  300
                                                                         
Db    179  ---THT-----------------------------FNRLFRR--------------------  188

Qy    301  PDYVEENRGHLIEIRDLFGPGLDTQEBPRIVILQQGAAAGIGKSTLARQVKEBAWGRGQLYGDR  360
                                                                          
Db    189  ---DEEGR-------------RPLTVVLQGPAGIGKTMAAKKILYDWAAGKLYQGQ  228

Qy    361  FQHVFYFSCRELAQSK--VVSLAELIGKPGTATPAPIRQLLSRPERLLFILDGYVDE-PGWV  418
                                                                         
Db    229  VDFAFFMPCGELLERPGTRSLADLILDQCPDRGAPVPQMLAQPQRLLFTILDGADELP--A  286

Qy    419  LQEPSSELCLIHWSQPQPADALLGSLLGKWTILPEASFLITARTTALQNLLIPSLEQARWEV  478
                                                                          
Db    287  LGGPEAAPCTDPFEAASGARVLGGLLSKALLPTALLLLVTTRAAAPGRLQGRLCSPQCAEV  346

Qy    479  LGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKEBLWALCLVPWWSWLACTCLMQQMKRKE  538
                                                                          
Db    347  RGFSDKDKKKYFYKFFRDERRAERAYRFVKGENETLFALCFVPFVCWIVCTVLRQQLELGR  406

Qy    539  KLTLTSKTTTTLCLHYLAQALQAQPL--GPQ---LRDLCSLAAEGIWQKKTLFSPDDLR  592
                                                                       
Db    407  DLSRTSKITTSVLLFITSVLSSAPVADGPRLQGDLRNLCRLAREGVLGRRAQPAEKELE  466

Qy    593  KHGLDGAIIST-FLKM---GILQEHPIPLSYSFIHLCFQEFFAAMSYVLED-------E  640
                                                                  
Db    467  QLEIRGSKVQTLFLSKKELPGVLETE---VTVQPIDQSFQEFLAALSVLLEDGGVPRTA  523

Qy    641  KGRG---------KHSNCIIDLEKTLEAYGIHGLPFGASTTRFLLGLLSDEGEREMENIFHC  692
                                                                          
Db    524  GGVGTLLRGDAQPHSHLVL---------TTRFLGLLSAERMRDIERHFGC  565

Qy    693  RLSQ----GRNLMQWV-------------PSLQLLLQP-----HSLESLHCL  722
                                                                
Db    566  MVSBRVKQEALRWVQGQGQGCPGVAPEVTEGAKGLEDTEEPEBEBEBEEGEEPNYPLELLYCL  625

Qy    723  YETRNKTFLTQVMAHFEEMGM------CVETDMELLVCTFCIKFSRHVKKLQLIEGROHRS  777
                                                                      
Db    626  YETQEDAFVRQALCRFPELALQRVRFC---RMDVAVLSCVCRCCPAGQALRLLSCR----  678

Qy    778  TWSPTMVVLFRWPVTDAYWQIILFSVLKYTRNI---KELDLSGNSLSHSAVKSLCKTLRR  834
                                                                    
Db    679  --------------LVAAQEKKKSLGKRLQASLGGGSSSQGTTKQLPASLLH  717

Qy    835  --------PRCLLETLRLAGCGLLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLC  886
                                                                      
Db    718  PLFQOAMTDPLCHLSSLITLSHCKLPDAVCRDLSEALRAAPALTELGLLNRLSEAGLRMLS  777

Qy    887  QRLRQPSCKLQ--RLQL--VSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLL  942
                                                                     
Db    778  EGLAWPQCRVQTVRVQLPDPQRGLI----QYLVGMLRQSPALTTLDLSGCQLPAPMVTYL  832

Qy    943  CEGLRHPACKLIRLGLDQTTLLSDEMRQSLRRALEQEKPQLLI  983
                                                   
Db    833  CAVLQHQGCGLQQTLSLASVELSEQSLQELQAVKRAKPDLVI  873
```

RESULT 11

MATE_MOUSE

```
ID   MATE_MOUSE              STANDARD;      PRT;   1111 AA.
AC   Q9R1M5; Q9JLR2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maternal antigen that embryos require (Mater protein) (Ooplasm-
DE   specific protein 1) (OP1).
GN   MATER.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIH Swiss; TISSUE=Ovary;
RX   MEDLINE=99360614; PubMed=10433232;
RA   Tong Z.-B., Nelson L.M.;
RT   "A mouse gene encoding an oocyte antigen associated with autoimmune
RT   premature ovarian failure.";
RL   Endocrinology 140:3720-3726(1999).
RN   [2]
RP   SEQUENCE FROM N.A. AND VARIANTS ASP-7; LEU-42; LYS-87; PRO-256;
RP   GLN-977 AND 1003-ASN-ASN-1004.
RC   STRAIN=129/Sv;
RX   MEDLINE=20222768; PubMed=10754103;
RA   Tong Z.-B., Nelson L.M., Dean J.;
RT   "Mater encodes a maternal protein in mice with a leucine-rich repeat
RT   domain homologous to porcine ribonuclease inhibitor.";
RL   Mamm. Genome 11:281-287(2000).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=20517328; PubMed=11062459;
RA   Tong Z.-B., Gold L., Pfeifer K.E., Dorward H., Lee E., Bondy C.A.,
RA   Dean J., Nelson L.M.;
RT   "Mater, a maternal effect gene required for early embryonic
RT   development in mice.";
RL   Nat. Genet. 26:267-268(2000).
CC   -!- FUNCTION: Necessary for embryonic development beyond the 2-cell
CC       stage.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Oocyte-specific.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF074018; AAD51762.1; -.
DR   EMBL; AF143573; AAF64393.1; -.
DR   EMBL; AF143559; AAF64393.1; JOINED.
DR   EMBL; AF143560; AAF64393.1; JOINED.
DR   EMBL; AF143561; AAF64393.1; JOINED.
DR   EMBL; AF143562; AAF64393.1; JOINED.
DR   EMBL; AF143563; AAF64393.1; JOINED.
DR   EMBL; AF143564; AAF64393.1; JOINED.
DR   EMBL; AF143565; AAF64393.1; JOINED.
DR   EMBL; AF143566; AAF64393.1; JOINED.
DR   EMBL; AF143567; AAF64393.1; JOINED.
DR   EMBL; AF143568; AAF64393.1; JOINED.
DR   EMBL; AF143569; AAF64393.1; JOINED.
DR   EMBL; AF143570; AAF64393.1; JOINED.
DR   EMBL; AF143571; AAF64393.1; JOINED.
DR   EMBL; AF143572; AAF64393.1; JOINED.
DR   PIR; A59000; A59000.
DR   HSSP; P10775; 2BNH.
DR   MGD; MGI:1345193; Mater.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   GO; GO:0009887; P:organogenesis; IMP.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; NACHT_NTPase.
DR   InterPro; IPR007111; NACHT_RNinh.
DR   Pfam; PF00560; LRR; 3.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat; Polymorphism.
FT   DOMAIN        1    128       5 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT        1     23       1 (INCOMPLETE).
FT   REPEAT       24     49       2.
FT   REPEAT       50     75       3.
FT   REPEAT       76    102       4.
FT   REPEAT      103    128       5.
FT   DOMAIN      191    513       NACHT.
```

```
Qy      275  NQKFTQLILLQRPHPRSQDPLVKRSWPDVVEENRGHLIRIDLFPGLDTQEPR---IVIL  332
             :|||  | |  |  | |:|     :         |     : | || |   |   |:|
Db      114  NKRFTKLLIAPGSGAGEDLLIGTSGEPEPERARRSDTHTFNRLF-RGNDDEGRPLTVVL  172

Qy      333  QGAAGIGKSTLARQVKEANGRGQLYGDRPQHVFYFSCRELAQSK--VVSLAELIGKDGYAT  391
             |    | ||:   |   :|| |   |  :|  :   |                : |:  |:
Db      173  QGPAGIGKTMAAKKILYDWAGGKLYHSQVDFAFFMPCGELLERPGTRSLLADLILEQCPDR  232

Qy      392  PAPIRQILSRPERLLFLFLDGVDE-PGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILP  450
             |||:|   || || :||| |:| | |||:|:   ::|    :|:|   | |:||| |  |:
Db      233  TAPVRRILAQPHRLLFILDGADELLP--TLAAPEATPCRDPFEATSGLRVLSGLLSQELLP  290

Qy      451  EASFLITARTTALQONLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAPRLVKSN  510
             :|:: |||| |  :|  :| | |:  :|: || |:|  ||    |:| ::|  || ||: |
Db      291  SARLLVTSRNATLGRLQGRLCSPQCAEVRGFSDKDKKKYFFKFPRDERKAERAYRPVKEN  350

Qy      511  KELWALCLVPWVSWLACTCLMQQMKRKEKLILTSKTTTTLCLHYLAQALQ-AQPLGP---  566
             : ||  |  |  :| |  | ||||:|:|  || |:|||||||:||| | | || ||
Db      351  ETLYALCFVPFVCWIVCTVLLQQMELGRDLSRISKITTTTSVYLLFITSMIKSAGTNGPRVQ  410

Qy      567  -QLRDLCSLAAEGIWQKKTLFSPDDLRKHIGLDGAIIST-FLKM----GILQBHIPLSYS  620
                |  |||   |  :|  :|| ||| |:|| |  :| |  | :     |||:|||| :|:
Db      411  GELRMLCRLAREGILKHQAQFSEKDLERLKLQSSQVQTMFLSKKELPGVLE---TVVIYQ  467

Qy      621  FIHLCFQEFFAAMSYVLEDEKGRGHKSNCIIDLEKTLEAYGIHGLFG--ASTTRFLLGLL  678
             |  |    |       ||  |   : | |   ::|:|  | :|:| |  ||| | | |||
Db      468  FIDQSFQEFLAALSYLLDAE---GAPGNSAGSVGMLINSDA--GLRGHLALTTRFLFGLL  522

Qy      679  SDEGEREMENIFHC-----RLSQGRNLMQWVPSIQLLLQP------------  713
             | :  :||:  |           ||::  |:| ||| | :|
Db      523  STERIRDIGNHFGCVVPGRVKQ--DTLRNV--QGQSQPKVATVGAEKKDELKDEEAEEE  577

Qy      759  KFSRHVKKLQLI--------------EGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNL  810
                  |  : ||:              |||:| |:|| | || ::| || | | |  | | |
Db      578  EEEEEEEBELNFGLELLVCLYETQEDDEVRQALSSLPE--MVLERVRLTRMDLEVLSYCV  635

Qy      811  KELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTLTEL  870
             ::|:| |||| |::|  | | | |  |:|| |   | |:| :| :|| |  |  ||:| |
Db      669  KGSQSTGKQPPASILRPLCEAMITQQCCGLSILTLSHCKLPDAVCRDLSEALKVAPSLREL  728

Qy      871  DLSFNVLIFDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQ  930
             | | || ||      | |:|||    | ::| ||| | | | ||  : | |||:|:|:|
Db      729  GLLQNRLTEAGLRLLSQGLAWPKCKVQTLRIQMPGL-QEVIHVLVITLQQSPVLTTLDLS  787

Qy      931  QNNLDDVGVRLLCEGLRHPACKLIRIGLDQTTLSDEMRQELRALEQEKPQLLIFSRRKPS  990
             |::||:|| | |  ||:  |||: |   |  | ::||:|| :    :| || || | |
Db      788  GCQLPGTVVEPLCSALKPKCGLKTLSLTSVELTENPLRELQAVKTLKPDLAIIHSKLGT  847

Qy      991  VMPTPTEG  997
             :| ||
Db      848  HPQPLKG  854

Qy      636  QCCPDGQPALRLVSCGLVAAKEKKKKKKSP-  668
                  | |: |:|:        ::   :::
Db      759  -------------------------------MNRL  668
```

```
FT   REPEAT        640    664       LRR 1.
FT   REPEAT        691    714       LRR 2.
FT   REPEAT        737    762       LRR 3.
FT   REPEAT        776    799       LRR 4.
FT   REPEAT        804    832       LRR 5.
FT   REPEAT        834    856       LRR 6.
FT   REPEAT        861    884       LRR 7.
FT   REPEAT        890    913       LRR 8.
FT   REPEAT        918    941       LRR 9.
FT   REPEAT        950    974       LRR 10.
FT   REPEAT        975   1003       LRR 11.
FT   REPEAT       1005   1027       LRR 12.
FT   REPEAT       1032   1059       LRR 13.
FT   NP_BIND       197    204       ATP (POTENTIAL).
FT   VARIANT      1003   1003       B -> D (IN STRAIN 129/SV).
FT   VARIANT         7      7       B -> D (IN STRAIN 129/SV).
FT   VARIANT        42     42       G -> D (IN STRAIN 129/SV).
FT   VARIANT        87     87       R -> K (IN STRAIN 129/SV).
FT   VARIANT       256    256       L -> P (IN STRAIN 129/SV).
FT   VARIANT       977    977       H -> Q (IN STRAIN 129/SV).
FT   VARIANT      1004   1004       SS -> NN (IN STRAIN 129/SV).
SQ   SEQUENCE     1111 AA;  125502 MW;  4016A5D67A1C01F4 CRC64;

Query Match             11.7%;  Score 882;  DB 1;  Length 1111;
Best Local Similarity   26.1%;  Pred. No. 1.2e-48;
Matches  274;  Conservative  154;  Mismatches  350;  Indels  272;  Gaps   25;
```

```
Qy    162 SSPDHESPSQOESPNAP-----TSTAVLGSWGSPPQPS----LAPR----EQEAPGTQWPLD 209
          :   :: :                                         ::::
Db     28 TSPENDSKSIOKDQGPEQOTSESTMG-------PPEKDSKAILKARGLEEKQSESTMSPSE 83

Qy    210 ETSGIYTEIREREREKSEKGRPPWAAVVGTPPQOAHTSLQPHHHPWEPSVRESLCSTWPW 269
          ::                                       ::: :
Db     84 NVSRAILKDSGSEEVBQASERK------MTSPENDSKSIQKDQGPEQOTSETLQS----- 133

Qy    270 KNEDFNQKFTQLILLLQRPHPRSQDPLVKRSWPDYVEENRGHLIE----------- 313
          : ::
Db    134 KEBD----EVTB--------------ADKDNGGD-LQDYKAHVIAKPDTSVDLHYDSPEMK 175

Qy    314 -IRDLFGPGLDTQEPRIVILQGAAGIGKSTLAROVKEAWGRGQLYGDRFQHVFYFSCREL 372
           : :      ::  : ::: :: :::::: :: :     :: : ::  :: :: ::: :
Db    176 LISDAFKPYQKTFQPHTILLHGRPGVGKSALARSTVLGWAQGKLF-QKMSFVIFFSVREI 234

Qy    373 AQSKVVSLAELIGKDGTAPTAPIRQILSRPERLLFILDGYDEPGWVLQEPSSELCLHWSQ 432
           ::  ::: ::        :  ::  ::::::::: :::: :   ::   :   :
Db    235 KWTEKSSLAQLIARKCPDSWDLVTKIMSQPERLLFVIDGLDDMDSVLHQDDMTLSRDWKD 294

Qy    433 PQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLGFSESSRKEYFYR 492
           :    :: :: :  :: ::  ::   :     : ::   ::  :::
Db    295 EQPIYILMYSLLRKALLPQSFLIITTRNTGLEKLKSMVVSPLYIVEGLSASRSQLVLE 354

Qy    493 YFFDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKKEKLTLTSKTTT--- 548
          :  : ::  ::   : :  :  ::      ::: ::  ::       :: :
Db    355 NISNESDRIQVFHSLIENHQLFDQCQAPSVCSLVCEALQLQKKLGKRCTLPCQTLTGLYA 414

Qy    549 TLCLHYL------AQALAQGPLGPQLRDLCSLAAAGIWQKKTLFSPDDLRKHGLDGAIIST 603
          ::                  ::    ::  : ::: ::     : :::
Db    415 TLVFHQLTLKRPSQSALSQEEQILVGLCMWAAGGVWTMRSVFYDDDLKNYSLKESEILA 474

Qy    604 FLKWGILQE--HPIPLSYSFIHLCPQEFFAAMSYVLEDEKGRGKH----SNCIIDLEK 655
          :   :: :           :: :  ::: :: :::                : :::
Db    475 LFHHNILLQVGHNSEQCYVFSHLSLQDFFAALYVLBGLEEWNQHFCFIENQRSIMEVKR 534

Qy    656 TLEAYGIHGLFGASTTRFLLGLLSDEGEREMENIFHCRL--SQGRNLMQWVPSLQLLLQP 713
          :           :::   ::: :                    : ::  :
Db    535 TDDT----RLLGMK--RFLFGLMNKDILKTLEVLFEYPVIPTVEQKLQHWV--SLIAQQ 585

Qy    714 ------HSLESLICLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKPSRHVKK 766
                :: :: :::                :     :     :::   :
Db    586 VNGTSPMDTLDAFYCLFESQDEEFVGGALKRFQEVWLLINQKMDLKVSSYCLKHCQNLKA 645

Qy    767 LQL----------IEGRQHRSTWSPTMVVL-FRWVPVTDAYWOILFSVLKVTRNLKELDLSGNSLS 821
          : :                                          :   ::::::: : ::::
Db    646 IRVDIRDLLSVDNTLELCPVVTVQSTQCKPLLMEWWGNPCSVLGSLRNLKELDLGDSILS 705
```

RESULT 12
PYA5_MOUSE

```
ID   PYA5_MOUSE    STANDARD;    PRT;    843 AA.
AC   Q91WS2; Q8K0L4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYRIN-containing APAF1-like protein 5-like.
GN   PYPAF5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   CONCEPTUAL TRANSLATION OF 1-287.
RA   Hinz U.;
RL   Unpublished observations (FEB-2003).
RN   [2]
RP   SEQUENCE OF 288-843 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA   Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT   "Identification of mammalian orthologs associated PYPAF5 with distinct
RT   functional roles.";
RL   FEBS Lett. 538:173-177(2003).
```

```
Qy    822 HSAVKSL------------ 828
          :: :
Db    706 QRAMKILCLELRNQSCRIQKLTFKSAEVVSGLKHLWKLLFSNQNLKYLNLGNTPMKDDM 765

Qy    829 ---CKTLRRPRCLLETLR-------------- 843
             : : :   ::  :: :
Db    766 KLACEALKHPKCCSVEFLRLDSCELTIIGYEMISTLLISTTRLKCLSLAKNRVGVKSMISL 825

Qy    844 -----------LAGCGLTAEDCKDLAFGLRANQTLIEDLSFNVLTDAGAKHLCQR 888
                       : :  :: ::   :   :  ::  :     :  : :
Db    826 GNALSSSMCLLQKLILDNCGLTPASCHLLVSALFSNQNLTHLCLSNNSLGTEGVQQLCQF 885

Qy    886 LRQPSCKLQR---------- 898
          :: :: ::
Db    889 LRNPECALQRLILNHCNIVDDAYGFLAMRLANNTKLTHLSLTMNPVGDAGAMKLLCEALKE 945

Qy    899 ---LQLVSCGLTSDCCODLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLHPAC 951
                 : ::     ::  ::  :   :::::: ::::    :: :::::
Db    946 PTCYLQELELVDCQLITQNCCEDLACMTTTTHKHLKSIDLGNNALGDKGVITLCEGLKQSS 1005

Qy    952 KLIRLGLDQTTLSDEMRQELRALEQEKPQL 981
          :: ::::      :   :: :
Db   1006 SLRRLGLGACKLITSNCCEALSIAISCNPHL 1035
```

```
CC   -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC       activation of NF-kappa-B (By similarity).
CC   -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC   -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC       sequences, in analogy to ortholog sequences.
CC   ------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------
DR   EMBL; BC013519; AAH13519.1; -.
DR   EMBL; BC031139; AAH31139.1; ALT_INIT.
DR   MGD; MGI:2141990; Pypaf5.
DR   InterPro; IPR007091; LRR.
DR   InterPro; IPR001611; LRR_RNinh.
DR   Pfam; PF00560; LRR; 1.
DR   PROSITE; PS50837; NACHT; 1.
DR   PROSITE; PS50824; DAPIN; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN       11    102       DAPIN.
FT   DOMAIN      168    485       NACHT.
FT   DOMAIN      569    585       ASP/GLU-RICH.
FT   DOMAIN      654    661       POLY-LYS.
FT   REPEAT      434    459       LRR 1.
FT   REPEAT      609    632       LRR 2.
FT   REPEAT      811    834       LRR 3.
FT   NP_BIND     174    181       ATP (POTENTIAL).
SQ   SEQUENCE   843 AA;  94592 MW;  35FB7A766A47DB51 CRC64;
```

```
Query Match     11.5%;  Score 868;  DB 1;  Length 843;
Best Local Similarity   31.6%;  Pred. No. 6.3e-48;
Matches 249;  Conservative 123;  Mismatches 281;  Indels 136;  Gaps 22;
```

```
Qy     275 NQKFTQLILLQRPHPRSQQDLVKRSWPDIVEENGHLIEIRDLI--GPGLDTQEPRIVILQ 333
                                                     :::::
Db     114 NKRFTKLLIAPGTGAVEDELLGPLGEPEPERARRSDTHTFNRLRFRGNDEBSSQPLTVVLQ 173

Qy     334 GAAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSK-VVSLAELIGKDGTATP 392
                                                    :::
Db     174 GPAGIGKTWAAKKILYDWAANGKLYHSQVDPAFFMPCGELLERPGKRSLADLVLDQCPDRA 233

Qy     393 APIRQILSRPERLLFILDGVDEPGWVLQEPSSEL---CLHWSQPQPADA-----LLGSLLG 445
                                                              :::::::::
Db     234 WPVKRILAQPNRLLFILDGADE---LPTLPSSEATPC------KDPLBATSGLRVLSGLLS 285

Qy     446 KTILPEASFLITAKTTALQNLIPSLEQARWVEVLGFSESSRKEYFYRYTFDERQAIRAFR 505
                                                               ::::
Db     286 QELLPGARLLVTTRHAATGRLQGRLCSPQCAEIRGFSDKDKCKYFFKFPRDERKAERAYR 345

Qy     506 LVKSNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTTLCLHYLAQALQ-AQPL 564
                                                          ::::::
Db     346 FVKENETLFALCFVPFVCWIVCTVLQQQLEIGRDLSRTSKTTTSVYLLFITSMLKSAGTN 405

Qy     565 GP----QLRLCSLAAEGIW-QKKTLFSPDDLRKHGLDGAIISTFLRKMGILQEHPIP--- 616
                                                          ::::
Db     406 GPRVQGELRTLCRLAREGILDHHKAQFSEEDLEKLKLRGSQVQIT--PLNKKEIPGVL 461

Qy     617 --LSYSFHILCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAYGIHGLFG------ 667
                                          :::::
Db     462 KTEVTYQFIDQSFQEFLAALSYLLEARTFG------------------TPAGGVQKLLNSDAELR 509

Qy     668 --ASTTRFLLFGLLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQLILQPH------ 714
                                          :::::
Db     510 GHIALTTRFLFGLLNTEGLRDIGNHIFGCVVPDHVYKKDTLRWVQG----QSHPKGPPVGA 564
```

```
RESULT 13
PYA6_HUMAN
ID   PYA6_HUMAN     STANDARD;      PRT;   1033 AA.
AC   P59045; Q8NBF5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYRIN-containing APAF1-like protein 6.
GN   PYPAF6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 365-1033 FROM N.A. (ISOFORM 2).
```

```
Qy     715 ------------------------------SLESLHCLYETRNKTFLITQVMAHFEEMGMCVE--- 746
                                         |||:||||:::||||||||:||||:||||| :
Db     565 KKTAELEDIEDAEEEEEEEEDLNFGLEBLLYCLYETQEEDFVRQALSSLPE--IVLERVRL 622

Qy     747 TDMELLVCTFCIKFSRHVKKLQLI-------------EGRQHRSTWSPTMVVLFRNWPVTDAYWQ 798
                ||:|||:|||||:|   |||        ||||:||::||:||||:||||:||||||
Db     623 TRMDLEVLANVCVQCCPDGQALRLVSCGLVAAKEKKKKKKS----------------- 662

Qy     799 ILFSVLKVTRNLKELDLSGNSLISHSSAVKSLCKTLRRPRCLLETLRLIAGCGLITABDCKDLA 858
                |||:| ::::||:|::|:||||:|:|||::|||:||||||||:||:::||||||:||||:
Db     663 ------------LVKRLKGSQSTKKQPPVSLLRPLCETMTPKCHLSVLIISHCRLPDAVCRDLS 715

Qy     859 FGLRANQTLTELDLSFNVLITDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVL 918
                :|||||::::||:|:||::|||||||||:|||||||||:|||||||:|||::||||||||
Db     716 EALKVAPALREIGLLQSRLTNTGLRLLCEGLAWPKCQVKTLRMQLPDL-QEVINILVIVL 774

Qy     919 SASPSLKELDLQQONNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEK 978
                ::::::||:|||:|:|||::||:||||||:||||||:|::|||::|||::|::|:||:||
Db     775 QOSPVITTLDDLSGCQLPGVIVEPLCAALRKHPCCSLKKTLSLITSVELSENSLRDLQAVKTSK 834

Qy     979 PQL-LIFSR 986
                ||: :|:|
Db     835 PDLSIIYSK 843
```

```
RC   TISSUE=Glial tumor;
RA   Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato T., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA   Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA   Yamashita H., Matsuo K., Nakamura Y., Saito K., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P59045-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P59045-2; Sequence=VSP_007068;
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC       Note=No experimental confirmation available;
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AY095145; AAM14632.1; -.
DR   EMBL; BC034730; AAH34730.1; -.
DR   EMBL; AK090621; BAC03490.1; ALT_INIT.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00560; LRR; 1.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00368; LRR_RI; 9.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
FT   DOMAIN        1     91       DAPIN.
FT   DOMAIN      147    470       NACHT.
FT   REPEAT      588    611       LRR 1.
FT   REPEAT      632    655       LRR 2.
FT   REPEAT      745    768       LRR 3.
FT   REPEAT      802    827       LRR 4.
FT   REPEAT      859    882       LRR 5.
FT   REPEAT      919    944       LRR 6.
FT   NP_BIND     153    160       ATP (POTENTIAL).
FT   VARSPLIC    614    668       SQMKSLVYWREICSLFYTMESLRELHIFDNDLNGISERILS
FT                                KALEHSSCKLRTLK -> R (in isoform 2).
FT                                /FTId=VSP_007068.
FT   CONFLICT    788    788       T -> I (IN REF. 3).
SQ   SEQUENCE   1033 AA;  117794 MW;  A28880485FBA49AC CRC64;
```

```
Qy       9.4%;
Query Match
Best Local Similarity    26.7%;                Score 711;  DB 1;   Length 1033;
Matches  229;  Conservative  126;   Mismatches 294;   Indels 208;   Gaps 20;
```

```
Qy    330  VILQGAAGTGKSTLARQVKEAWGRGQLYGDRFQHYFYFSCRELAQSKVVSLAELLIGKDGT 389
           ::|:|:|||  |: :|||::|||:    |  :   |||||:|||     | ||:|:
Db    149  VFLMGERASGKTIVINLAVLRWIKGEMWQNMISYVVHLTSHEINQMTNSSLAELIIAKDWP 208

Qy    390  ATPAPIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTIL 449
           || |::||:||:||:|||:||:    ||:| |   || :|||||   | |  |  |  |
Db    209  DQQAPIADILSDPKLLFILBDLIDNIRFELNVNESALICSNSTQKVPIPVLLIVSILKRMA 268

Qy    450  PEASFLITARTTALQNLIPSLEQARWEVLGFSESSSRKEYFYRYFIFTDERQAIRAFRLVKS 509
           |:|||  :  |: |::|  :  || |  |  |  :|  |::|  |: |:|:||::|: :
Db    269  PGCWFLISSRPTRGNNVKTFLKEVDCCTTLQLSNGKREIYFNSFFKDRQRASAALQLVHE 328

Qy    510  NKELWALCLYPWVSWLACTCLMQQMRKEKLTLISKTTTLCLHYLAQALQAQP------ 563
           |||:|  || | :| |  |   |  ||:||  |:| :| |||:|  |||  ||
Db    510  ... 563
```

```
RESULT 14
ID   CARD8_HUMAN       STANDARD;       PRT;       431 AA.
AC   Q9Y2G2; Q96P82;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Caspase recruitment domain protein 8 (Apoptotic protein NDPP1)
DE   (CARD-inhibitor of NF-kappaB activating ligand) (CARDINAL) (TUCAN)
DE   (CARD8 OR NDPP1 OR KIAA0955.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
```

```
Qy    329  DEILVGLLCRVAILCWTTCTVULKRQMDKGRDFQLCCQQTTPDLHAAHFLADALITSEAGLTANQ 388
           ::|| :  |||   ||| :|  ||:||      | :  | :|::| || :| : :
Db    329  -LGPQLRDICLSLAAEGTWQKKTLPSPDDLRKHGLDGAIISTFLKMGILQEHIPLS--- 618

Qy    389  YHLG-LLKRLCLLAAGGLFLSTLNFSGEDLRCVGFTEADVSVLQOAANIL----LPSNTHK 443
           |: | || || |   | |:|  ::| :  | ||   | :| :|:  |:::      :::
Db    389  YHLG-LLKRLCLLAAGGLFLSTLNFSGEDLRCVGFTEADVSVLQOAANIL----LPSNTHK 443

Qy    444  DRYKFIHLNVQEFCTAIAFLMAVPNYLIPSGSREYKEKRREQYSD------ 487
           || ||  || ||| | |||:| |||| |||||| ||| |  ||:
Db    444  --YSFIHLCRQFFFAAMSYVL----- -EDEKGRGKHSNCIIDLEKTLEAYGIHG 664

Qy    488  -FNQVFT-FIFGLLNANRKILETSFGYQLPMVDSFKNYSVGYMKHLDRDPEKLTHHMPL 545
           | :||  |:||:| :|:| :||: | |  || :: :  |:  | | | :|   |
Db    488  ... 545

Qy    546  FYCLYENREBEFVKRTIVDALMEVTTVVLQSDKDOMVSLYCLDYCCHLLRTLKLSVQRIFQNK 605
           : ||||||| |||:  | :|| |  |  | || | :  | | |: | || |  : :|:
Db    546  ... 605

Qy    606  EPLIRPPASQMKSLV----YWREICSLFYTMESLRELHIFDNDLNGISERILSKALEHS 660
           ||| |   |:|:|     |:||:|||||||||||||||||||||||||||||||||||
Db    606  ... 660

Qy    661  SCKLRTLKLSYVSTASGFEDLLKALARNRSLTIYLSINCTSISLNMFSLLHDILHEPTCQI 720
           |||||||||| :|| | ||:| || | | || | |:| |:| | |  :|::|:| |  :|
Db    661  ... 720

Qy    721  SHLSLMKCDLRASECEBIASLLISGGSLRKLTLSSNPLRSDGWNILCDALLHPNCTLISL 780
           || |:| ||| | | :||:| |||||||||||||||| || ||:||||||:|||| |||
Db    721  ... 780

Qy    781  VLVFCCLTENCCSALGRVLLFSPTLRQLDLCVNRLKNYGVLHVTFPLLFPTCQLEELHLS 840
           || |||| |||| ||| ||: |: || || |  | :| | :  |:| ||| :|:|| |
Db    781  ... 840

Qy    824  AVK----------SLCKTLRR---------------------PRCLL 839
           ::|          |:| ||::                     |:||
Db    824  ... 839

Qy    840  ETLRLAGCGLLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRL 899
           ||||| | ||| ||||| |  |||||| |:|||||||||| | | ||| :|:| ||||:|
Db    840  ... 899

Qy    900  QLVSCGLTSDCCQDLASVLSASPSLSKELDL------------------ASVLITNKTLE 929
           |:| |||| |||||||| |   ||||| ||                   |||||:| ||
Db    900  ... 929

Qy    901  LTTSACCRSL------------------ASVLITNKTLE----------RLNLLQNHLGNDGVA- 935
           :| : |||||                  :|||:|  ||          |||||:| | | |
Db    901  ... 935

Qy    930  ---------------QONNLDDVGVRLLCEGLRHPACKLIRLGLDQTT 962
           |:|:||||||:|||||:||||:|||:||||
Db    936  ... 944

Qy    963  LSDEMRQELRALEQEKPQLLIFPSRRKPSVMPTEGLDTGEMSNSTSSLKRQRLGSERAAS 1022
           || :| |:| :||  | |  | ::  :|  |:|||  :| |  |:|| |: |:| | |
Db    963  ...

Qy    1023 HVAQANLKLLDVSKIFP 1039
           |: ||| || |  |:|
Db    1023 KLLE-SLISP 944
```

RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=21950691; PubMed=11821383;
RA   Razmara M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
RA   DiStefano P.S., Bertin J., Alnemri E.S.;
RT   "CARD-8 protein, a new CARD family member that regulates caspase-1
RT   activation and apoptosis.";
RL   J. Biol. Chem. 277:13952-13958(2002).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RA   Zhang H.;
RT   "A novel apoptotic protein, NDPP1, containing CARD and BH3 domains.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RA   Guiet C., Vito P.;
RT   "DACAR, a novel CARD-containing protein.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=21570185; PubMed=11551959;
RA   Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
RA   MacFarlane M., Martin S.J.;
RT   "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
RT   of multiple NF-kappa B activation pathways.";
RL   J. Biol. Chem. 276:44069-44077(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Kidney;
RA   Guo J.H., Yu L.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   CHARACTERIZATION.
RA   Pathan N., Marusawa H., Krajewska M., Matsuzawa S.-I., Kim H.,
RA   Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Pio P.,
RA   Godzik A., Reed J.C.;
RT   "TUCAN, an antiapoptotic caspase-associated recruitment domain family
RT   protein overexpressed in cancer.";
RL   FEBS Lett. 521:165-169(2002).
RN   [8]
RP   CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
RX   MEDLINE=22062958; PubMed=12067710;
RA   Stilo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
RT   "TUCAN/CARDINAL and DRAL participate in a common pathway for
RT   modulation of NF-kappaB activation.";
RL   J. Biol. Chem. 276:33220-33229(2001).
CC   -!- FUNCTION: inhibits NF-kappaB activation. May participate in a
CC   regulatory mechanism that coordinates cellular responses
CC   controlled by NF-kappaB transcription factor. Involved in the
CC   negative regulation of caspase-1.
CC   -!- SUBUNIT: May form homodimers. Interacts with NEMO and DRAL. Binds
CC   to caspase-1, pseudo-ICE and ICEBERG.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=Long;
CC   IsoId=Q9Y2G2-1; Sequence=Displayed;
CC   Name=Short;
CC   IsoId=Q9Y2G2-2; Sequence=VSP_000782, VSP_000783;
CC   -!- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
CC   placenta. Lower expression in heart, kidney and liver. Also
CC   expressed in spleen, lymph node and bone marrow.
CC   -!- SIMILARITY: Contains 1 CARD domain.
DR   EMBL; AB023172; BAA76799.1; -.

DR   EMBL; AF322184; AAG50014.1; -.
DR   EMBL; AF331519; AAK01126.1; -.
DR   EMBL; AY026322; AAK08982.1; -.
DR   EMBL; AF405558; AAL02427.1; -.
DR   EMBL; AF511652; AAM46959.1; -.
DR   InterPro; IPR001315; CARD.
DR   Pfam; PF00619; CARD; 1.
DR   SMART; SM00114; CARD; 1.
DR   PROSITE; PS50209; CARD; 1.
KW   Apoptosis; Nuclear protein; Alternative splicing.
FT   DOMAIN   340   430   CARD.
FT   VARSPLIC   287   431   Missing (in isoform Short).
FT   /FTId=VSP_000782.
FT   VARSPLIC   282   286   ELKLS -> WISSL (in isoform Short).
FT   /FTId=VSP_000783.
FT   MUTAGEN   366   366   L->R: INHIBITS HOMODIMER FORMATION.
FT   CONFLICT   60   60   E -> G (IN REF. 5).
FT   CONFLICT   326   326   V -> M (IN REF. 5).
FT   CONFLICT   422   422   L -> P (IN REF. 5).
SQ   SEQUENCE   431 AA; 48932 MW; CB54D13080773286 CRC64;

Query Match   Best Local Similarity   39.5%;
                        Score 668; DB 1; Length 431; Pred. No. 1.7e-35;
Matches 151; Conservative 53; Mismatches 122; Indels 56; Gaps 7;

Qy   1079 FWGPTGPVATEVVDKEKNLLRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLG- 1137
Db   56 FLGPEGNVDVELIDKSTNRYSVWPPTAGWYLWSATGLGFLVRDEVTVTIARGSWSQHLAL   115

Qy   1138 EINPQHSWMVAGPLLDIKAEP-GAVEAVHLPHFVALQGGHVDTSLFQMAIFKEEGMLLEK 1196
Db   116 DLQHHEBQWLVGGGPLPDVTAEPEEAVAEIHLPHFISLQ-GEVDVSNFLVAIFKNEGGNVLEH 174

Qy   1197 PARVELHHIVLENPSPSPsPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSD 1256
Db   175 PARVEPFYAVLESPSFSLMGILLRIASGTRLSIPITSNTLIYHPEDIKFHLYLVPSD   234

Qy   1257 CSIR----------------------------------KELELCYRSPGEDQ 1274
Db   235 ALLTKAIDDEEDRFHGVRLQTSPPMEPLNFGSSYIVSNSSANLKVMPKELKLSYRSPGEIQ   294

Qy   1275 LFSREFYVGHLGSGIRLQVRDKDKCDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQL 1334
Db   295 HFSKPYAGQMKEPIQLEITEKRHGTLVWDTEVKPVDLQLVAASAPP-------PFSGAA-   346

Qy   1335 LHFVDQYRELQIARVTSVEVVLDKLH-GQVLSQEQYERVLAENTRPSQMRKLFSLSQSWD 1393
Db   347 --FVKENHRQLQARMGDLKGVLDDLQDNEVLTENEKELVEQEKTRQSKNEALLSMVEKKG   404

Qy   1394 RKCKDGLIYQALKETHPHLIMEL 1415
Db   405 DLALDVLFRSISERDPYLVSYL 426

RESULT 15
CARP_HUMAN
ID   CARP_HUMAN   STANDARD;   PRT; 1040 AA.
AC   Q9H029; Q96RH5; Q96RH6; Q96RH8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory
DE   bowel disease protein 1).
GN   CARD15 OR NOD2 OR IBD1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND
RP   VARIANT ARG-908.
RC   TISSUE=Breast;
RX   PubMed=11087742;

```
RA   Ogura Y., Inohara N., Benito A., Chen F.F., Yamaoka S., Nunez G.;
RT   "Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and
RT   activates NF-kappaB.";
RL   J. Biol. Chem. 276:4812-4818(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS CD THR-140, ARG-157,
RP   CYS-235, ARG-248, SER-268, SER-289, ASN-291, VAL-301, TRP-311,
RP   VAL-348, CYS-373, SER-414, LEU-431, VAL-432, LYS-441,
RP   VAL-612, THR-612, TRP-684, TRP-702, CYS-703, CYS-713, GLY-725,
RP   VAL-755, VAL-758, LYS-778, MET-793, LYS-843, SER-853, VAL-863,
RP   ARG-908, ASP-918, ASP-924 AND ILE-955, VARIANTS ULCERATIVE COLITIS
RP   THR-140 AND THR-885, AND VARIANT MET-189.
RC   TISSUE=Leukocyte;
RX   MEDLINE=21279172; PubMed=11385576;
RA   Hugot J.-P., Chamaillard M., Zouali H., Lesage S., Cezard J.-P.,
RA   Belaiche J., Almer S., Tysk C., O'Morain C.A., Gassull M., Binder V.,
RA   Finkel Y., Cortot A., Modigliani R., Laurent-Puig P.,
RA   Gower-Rousseau C., Macry J., Colombel J.-F., Sahbatou M., Thomas G.;
RT   "Association of NOD2 leucine-rich repeat variants with susceptibility
RT   to Crohn's disease.";
RL   Nature 411:599-603(2001).
RN   [3]
RP   VARIANTS BS GLN-334, TRP-334 AND PHE-469.
RX   MEDLINE=21419644; PubMed=11528384;
RA   Miceli-Richard C., Lesage S., Rybojad M., Prieur A.M.,
RA   Manouvrier-Hanu S., Hafner R., Chamaillard M., Zouali H., Thomas G.,
RA   Hugot J.-P.;
RT   "CARD15 mutations in Blau syndrome.";
RL   Nat. Genet. 29:19-20(2001).
CC   -!- FUNCTION: Induces NF-kappaB via RICK (CARDIAK, RIP2) and IKK-
CC       gamma. Confers responsiveness to intracellular bacterial
CC       lipopolysaccharides (LPS).
CC   -!- SUBUNIT: Binds to RICK by CARD-CARD interaction.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, 1/Nod2 (shown here) and 2/Nod2b, are
CC         produced by alternative initiation. Both isoforms can activate
CC         NF-kappaB. Isoform 1 is more abundant;
CC   -!- TISSUE SPECIFICITY: Monocytes-specific.
CC   -!- DISEASE: Defects in CARD15 are the cause of Blau syndrome (BS)
CC       [MIM:186580], a rare autosomal dominant disorder characterized by
CC       early-onset granulomatous arthritis, uveitis and skin rash.
CC   -!- DISEASE: Defects in CARD15 are a cause of susceptibility to
CC       Crohn's disease (CD) [MIM:266600], a form of remitting
CC       inflammatory bowel disease. CD may involve any part of the
CC       gastrointestinal tract, but most frequently the terminal ileum and
CC       colon. Bowel inflammation is transmural and discontinuous. Crohn's
CC       disease is commonly classified as autoimmune disease.
CC   -!- DISEASE: Defects in CARD15 are a cause of susceptibility to
CC       ulcerative colitis [MIM:191390], a chronic inflammatory bowel
CC       disease. In ulcerative colitis, the inflammation is continuous and
CC       limited to rectal and colonic mucosal layers. Ulcerative colitis
CC       is commonly classified as autoimmune disease.
CC   -!- SIMILARITY: Contains 2 CARD domains.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
```

```
DR   EMBL; AF178930; AAG33677.1; -.
DR   EMBL; AF385009; AAK70867.1; -.
DR   EMBL; AF385090; AAK70868.1; -.
DR   EMBL; AJ303140; CAC42117.1; -.
DR   Genew; HGNC:5331; CARD15.
DR   MIM; 605956; -.
DR   MIM; 186580; -.
DR   MIM; 266600; -.
DR   MIM; 191190; -.
DR   InterPro; IPR003315; CARD.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007109; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   Pfam; PF00560; LRR; 2.
DR   SMART; SM00114; CARD; 1.
DR   PROSITE; PS50209; CARD; 2.
DR   PROSITE; PS50837; NACHT; 1.
DR   PROSITE; PS50007; NACHT; 1.
KW   ATP-binding; Repeat; Leucine-rich repeat; Disease mutation;
KW   Alternative initiation; Polymorphism.
FT   CHAIN         1   1040   CASPASE RECRUITMENT DOMAIN PROTEIN 15,
FT                            ISOFORM 1.
FT   CHAIN        28   1040   CASPASE RECRUITMENT DOMAIN PROTEIN 15,
FT                            ISOFORM 2.
FT   INIT_MET     28     28   FOR ISOFORM 2.
FT   DOMAIN       26    122   CARD 1.
FT   DOMAIN      126    218   CARD 2.
FT   NP_BIND     299    306   ATP (POTENTIAL).
FT   DOMAIN      293    618   NACHT.
FT   REPEAT      786    812   LRR 1.
FT   REPEAT      814    837   LRR 2.
FT   REPEAT      842    865   LRR 3.
FT   REPEAT      926    949   LRR 4.
FT   REPEAT      954    977   LRR 5.
FT   REPEAT      982   1005   LRR 6.
FT   VARIANT     140    140   A -> T (in CD and ulcerative colitis).
FT                            /FTId=VAR_012665.
FT   VARIANT     157    157   W -> R (in CD).
FT                            /FTId=VAR_012666.
FT   VARIANT     189    189   P -> S.
FT                            /FTId=VAR_012667.
FT   VARIANT     235    235   N -> S
FT                            /FTId=VAR_012668.
FT   VARIANT     248    248   R -> C (in CD).
FT                            /FTId=VAR_012669.
FT   VARIANT     268    268   L -> R (in CD).
FT                            /FTId=VAR_012670.
FT   VARIANT     289    289   T -> M.
FT                            /FTId=VAR_012671.
FT   VARIANT     291    291   D -> N (in CD).
FT                            /FTId=VAR_012672.
FT   VARIANT     294    294   T -> S (in CD).
FT                            /FTId=VAR_012673.
FT   VARIANT     301    301   A -> V (in CD).
FT                            /FTId=VAR_012674.
FT   VARIANT     311    311   R -> W (in CD and ulcerative colitis).
FT                            /FTId=VAR_012675.
FT   VARIANT     334    334   R -> Q (in BS).
FT                            /FTId=VAR_012676.
FT   VARIANT     334    334   R -> W (in BS).
FT                            /FTId=VAR_012677.
FT   VARIANT     348    348   L -> V (in CD).
FT                            /FTId=VAR_012678.
FT   VARIANT     352    352   H -> R (in CD).
FT                            /FTId=VAR_012679.
FT   VARIANT     373    373   R -> C (in CD).
FT                            /FTId=VAR_012680.
FT   VARIANT     414    414   N -> S (in CD).
FT                            /FTId=VAR_012681.
FT   VARIANT     431    431   S -> L (in CD).
FT                            /FTId=VAR_012682.
FT   VARIANT     432    432   A -> V (in CD).
FT                            /FTId=VAR_012683.
FT   VARIANT     441    441   E -> K (in CD).
FT                            /FTId=VAR_012684.
FT   VARIANT     469    469   L -> F (in BS).
FT                            /FTId=VAR_012685.
FT   VARIANT     612    612   A -> T (in CD).
FT                            /FTId=VAR_012686.
FT   VARIANT     612    612   A -> V (in CD).
FT                            /FTId=VAR_012687.
```

```
FT   VARIANT       684    684    R -> W (in CD).
FT                                /FTId=VAR_012688.
FT   VARIANT       702    702    R -> W (risk factor for CD).
FT                                /FTId=VAR_012689.
FT   VARIANT       703    703    R -> C (in CD and ulcerative colitis).
FT                                /FTId=VAR_012690.
FT   VARIANT       713    713    R -> C (in CD).
FT                                /FTId=VAR_012691.
FT   VARIANT       725    725    A -> G (in CD).
FT                                /FTId=VAR_012692.
FT   VARIANT       755    755    A -> V (in CD and ulcerative colitis).
FT                                /FTId=VAR_012693.
FT   VARIANT       758    758    A -> V (in CD).
FT                                /FTId=VAR_012694.
FT   VARIANT       778    778    E -> K (in CD).
FT                                /FTId=VAR_012695.
FT   VARIANT       793    793    V -> M (in CD).
FT                                /FTId=VAR_012696.
FT   VARIANT       843    843    E -> K (in CD).
FT                                /FTId=VAR_012697.
FT   VARIANT       853    853    N -> S (in CD).
FT                                /FTId=VAR_012698.
FT   VARIANT       863    863    M -> V (in CD).
FT                                /FTId=VAR_012699.
FT   VARIANT       885    885    A -> T (in ulcerative colitis).
FT                                /FTId=VAR_012700.
FT   VARIANT       908    908    G -> R (in CD).
FT                                /FTId=VAR_012701.
FT   VARIANT       918    918    A -> D (risk factor for CD).
FT                                /FTId=VAR_012702.
FT   VARIANT       924    924    G -> D (in CD).
FT                                /FTId=VAR_012703.
FT   VARIANT       955    955    V -> I.
FT                                /FTId=VAR_012704.
FT   MUTAGEN       305    305    K->R: NO ACTIVATION.
SQ   SEQUENCE     1040 AA;  115282 MW;  0037592D96D7DFF CRC64;

Query Match             5.2%;
Best Local Similarity   25.4%;
Matches  198;  Conservative  111;  Mismatches  325;  Indels  147;  Gaps   31;

Score 393.5;  DB 1;  Length 1040;
Pred. No. 2.8e-17;

Qy    314  IRDLFG-PGLDTQEPRIVILQGRAAGIGKSTLARQVKEAWGRGQLYGDRPQHVFYSCREL  372
                :  : | ||| :: || | |||:||| |::   ::
Db    278  LEELFSTPGHLNDDADTVLVVGEAGSGKSTLLQRLHLLWAAGQDF-QEFLFVFFPSCRQL  336

Qy    373  -AQSKVVSLAEL----------IGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQBP  422
               :  |:|:|                :||:||| :||: | |||::
Db    337  QCMAKPLSVRTLLFEHCCWPDVQGEDI-----FQLILDHPDRVLLTPDGPDE--FKFRPT  389

Qy    423  SSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFS  482
           |:||  |  : ||| ||||:||||||| |:||: ||| :|  | ||:|  | ||||||
Db    390  DRERHCSPFDPTSVQTLLFNLLQGNLLKNARKVVTSRPAAVSAFLRKYIRTEF-NLKGFS  448

Qy    483  ESSRKEYFYRYFTDBRQAIRAFRLVKSNKBLWALCLVPWVSWLACTCLMQQMRKEKLTL  542
           |:  ::|:  || : ||| | |: || ::| | |  |: :: :| ||:||
Db    449  EQGIELYLRKRHHEPGVADRLIRLLQETISALHGILCHLFVPSWMVSKC------HQELL  501

Qy    543  ----TSKTTTTLCL------HYLAQAL---QAQPLGLP-----QLRDLCSLAAEGIW----  580
               ||||||||:|      ||||:|    |||:| |      ||::|: | :|
Db    502  QEGGSPKITTTDMYLLLQHFLILHATPPDSASQGLGPSLLRGRLPTLLHLGRLALWGLGMC  561

Qy    581  ------QKKTLFSPDDLRKHGLDGAIISTFLXMGILLQEHPIPLSYSPIHLCFQEFFAA  632
                 | :|:|||| :|   ||: | ::| : |:    |::  |:|: |||:||||
Db    562  CYVFSAQQLQAAQVSPDDISL----GFLVRA---KGVVPGSTAPL--EFLHITFQCFPAA  612

Qy    633  MSVYLBD-------------------EKGRGKHSNCIIDLBKTLBAYGI  662
           ||: :                       ||| :| | :|::||:  ||
Db    613  FYLALSADVPPALLRHLFNCGRPGNSPMARLLPTMCIQASEGKDSSVAALLQKA-EPHNL  671

Qy    663  HGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNL-----MQWVPSLQLLLQPHSL-  716
           | |||:|:   |||:|| :       ||   ||     :    ||| : | | ||||:
Db    672  Q------ITAAFLAGLLSRE----HWGLLAECQTSEKALLRRQACARWCLARSLRHFHSTP  723

Qy    717  ------ESLHC-------------LYETRNKTFLTQV------MAHFEBMGMCVBTDMELLVC  754
                 : | :| |   :|: :  ::: :: :| : :| | | |:|:|
Db    724  PAAPGEAKSVHAMPGFIWLIRSLYEMQEERLARKAARGLNVGHL-KLTPCSVGPVTECAAL  782

Qy    755  TFCIKFSRHVKKLQL-------IEGRQHRSTWSPIMVVLPRWVPVTDAYWQILFSVLKVT  807
             |::| : :: | |       | ||  :  |||::|:| | | || |:: | ||:|:
Db    783  APVLQHLRRPVALQLDYNSVGDIGVEQLLPCLGVCKALYLRDNNISDRGICKLIECALHC  842

Qy    808  RNIKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLACGLTABDCKDLAFGLRANQTL  867
           .: |:|:| || ||:::|| | ||: | || | || ||| |||:|:||| | ||||||
Db    843  EQLQKLALFNNKLTDGCAHSMAKLLACRQNFL-ALRLGNNYITAAGAQVLABGLRGNTSL  901

Qy    868  TELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKEL  927
            ||: | ::| | | |:|| ::|: |  | ||  :|| | |: |:||| ::|:| ::|||
Db    902  QFLGFWGNRVGDEGAQALABEALGDHQ-SLRWLSLVGNNIGSVGAQALALMLAKNVMLEEL  960

Qy    928  DLQQNNLDDVGVRLLCEGL-RHPACKLIRLGLDQTT-LSDEMRQELRALLEQEKPQLLIFS  985
            |:::|| | || ||| || :::  ||| ||:: :: | || :||| || ::|::|||:
Db    961  CLEENHLQDBGVCSLABGLKKNNSLKILIKLSNNCITYLGABEAL--LQALERNDTILEWL  1018

Qy    986  R  986
           |
Db   1019  R  1019

Search completed: January 29, 2004, 13:46:30
Job time : 26.9982 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 79.1934 Seconds
                  4656.416 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWGRLACYLFPLKKEE.........HLIMELWEKGSKKGLLPLSS 1429

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5
Searched: 830525 seqs, 258052604 residues
                  (without alignments)
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   SPTREMBL.23:*

```
1:  sp_archea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacterial:*
17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	2640	35.0	509	4	Q9H5Z8	Q9h5z8 homo sapien
2	1754	23.3	409	4	Q9H5Z7	Q9h5z7 homo sapien
3	1371.5	18.2	892	4	Q8NEU4	Q8neu4 homo sapien
4	1254.5	16.7	1033	11	Q8R4B8	Q8r4b8 mus musculu
5	899.5	11.9	982	11	Q8BU40	Q8bu40 mus musculu
6	868	11.5	863	11	Q8C6J9	Q8c6j9 mus musculu
7	843	11.2	846	4	Q8IXT0	Q8ixt0 homo sapien
8	815.5	10.8	713	6	Q95LZ7	Q95lz7 macaca fasc
9	689.5	9.2	748	11	Q99MW0	Q99mw0 mus musculu
10	679.5	9.0	825	11	Q8C6M5	Q8c6m5 mus musculu
11	676	9.0	657	11	Q9EPG7	Q9epg7 mus musculu
12	573	7.6	673	11	Q8CCN1	Q8ccn1 mus musculu
13	393.5	5.2	953	4	Q8IWF5	Q8iwf5 homo sapien
14	392	5.2	1020	11	Q8K3Z0	Q8k3z0 mus musculu
15	383	5.1	287	11	Q9BY26	Q9by26 homo sapien
16	372.5	4.9	778	4	Q8NF48	Q8nf48 homo sapien
17	363.5	4.8	456	11	Q91VI7	Q91vi7 mus musculu
18	361.5	4.8	456	11	Q92P4	Q92p4 mus musculu
19	358	4.8	447	4	Q96FD7	Q96fd7 homo sapien
20	358	4.8	461	4	Q9BQ80	Q9bq80 homo sapien
21	357	4.7	461	6	Q8HZP9	Q8hzp9 pan troglod
22	351.5	4.7	461	11	Q8BW21	Q8bw21 mus musculu
23	351	4.7	390	11	Q8BHB0	Q8bhb0 mus musculu
24	348	4.6	461	4	Q8IZK8	Q8izk8 homo sapien
25	339.5	4.5	516	4	Q8NBF5	Q8nbf5 homo sapien
26	334	4.4	1052	7	Q9GJD8	Q9gjd8 rattus norv
27	334	4.4	1073	7	Q9GJD9	Q9gjd9 rattus norv
28	334	4.4	1153	7	Q9GJE0	Q9gje0 rattus norv
29	325	4.3	269	11	Q9D458	Q9d458 mus musculu
30	318	4.2	1155	7	Q9TPP1	Q9tpp1 mus musculu
31	316.5	4.2	519	11	Q8C249	Q8c249 mus musculu
32	304	4.0	660	4	Q8NF06	Q8nf06 homo sapien
33	289	3.9	733	4	Q8TEE2	Q8tee2 homo sapien
34	289	3.8	977	7	Q8HW99	Q8hw99 mus musculu
35	280.5	3.7	884	7	Q29675	Q29675 homo sapien
36	278.5	3.7	932	4	Q96KL4	Q96kl4 homo sapien
37	277	3.7	1056	4	Q96K14	Q96k14 homo sapien
38	276	3.7	692	4	Q8NF42	Q8nf42 homo sapien
39	272.5	3.6	706	11	Q96D51	Q96d51 homo sapien
40	258	3.4	1097	4	Q8BUT6	Q8but6 mus musculu
41	241	3.2	195	11	Q8HXK9	Q8hxk9 homo sapien
42	232.5	3.1	1402	11	Q9H6Y0	Q9h6y0 homo sapien
43	232.5	3.1	1403	11	Q8C365	Q8c365 mus musculu
44	226.5	3.0	1402	11	Q8CH70	Q8ch70 mus musculu
45	223.5	3.0	1403	11	Q8CG74	Q8cg74 mus musculu
				11	Q8BG68	Q8bg68 mus musculu

RESULT 1

Q9H5Z8

```
ID   Q9H5Z8      PRELIMINARY;    PRT;    509 AA.
AC   Q9H5Z8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ22740.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA   Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA   Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA   Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA   Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK026393; BAB15469.1; -.
DR   HSSP; P10775; 2BNH.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   Pfam; PF00560; LRR; 1.
DR   PROSITE; PS50103; LRR_RI; 2.
DR   PROSITE; PS50050; LRR; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   509 AA;  56908 MW;  EB7535AF69817F5B CRC64;
```

ALIGNMENTS

Query Match   35.0%; Score 2640; DB 4; Length 509;
Best Local Similarity   100.0%;
Matches 506; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy   735 MAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTWVVLFRWVPVTD 794
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTWVVLFRWVPVTD 60

Qy   795 AYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDC 854
```

```
Db   61 AVWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGGGLTABDC 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  855 KDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDL 914
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 KDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDL 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  915 ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRAL 974
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRAL 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  975 EQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHVAQANLKLLDV 1034
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHVAQANLKLLDV 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1035 SKIFPIAEIAEHSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGPTGPVATEVVDKE 1094
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 SKIFPIAEIAEHSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGPTGPVATEVVDKE 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1095 KNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEINPQHSWMWVAGPLLDI 1154
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 KNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEINPQHSWMWVAGPLLDI 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1155 KAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARVEL 1214
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARVELHHIVLENPSFSP 480
        |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1215 LGVLLKMIHNALRFIPVTSVVLLYHR 1240
        ||||||||||||||||||||||||||
Db  481 LGVLLKMIHNALRFIPVTSVVLLYHR 506
```

RESULT 2
```
Q9H5Z7

ID   Q9H5Z7               PRELIMINARY;      PRT;   409 AA.
AC   Q9H5Z7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ22745.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA   Nakajima Y., Mizuno Y., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA   Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA   Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA   Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK026398; BAB15470.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   409 AA;  46068 MW;  07CC5FACF3EB7236 CRC64;
```

```
Query Match             23.3%;  Score 1754;  DB 4;  Length 409;
Best Local Similarity   87.7%;  Pred. No. 1.1e-123;
      Matches  343;  Conservative   0;  Mismatches   0;  Indels   48;  Gaps   2;

Qy  967 MRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHVAQ 1026
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MRQELRALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHVAQ 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1027 ANLKLLDVSKIFPIAEIA-----EESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGP 1082
        |||||||||||||||||     ||||||||||||||||||||||||||||||||||||||
Db   61 ANLKLLDVSKIFPIAEIAGKSHEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGP 120
        ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Qy 1083 TGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEINPQ 1142
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEINPQ 180
```

RESULT 3
```
Q8NEU4

ID   Q8NEU4               PRELIMINARY;      PRT;   892 AA.
AC   Q8NEU4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Monarch-1 splice form IV.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Williams K.L., Linhoff M.W., Ting J.P.Y.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL   EMBL; AY116207; AAM75145.1; -.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50503; LRR_RI; 1.
DR   PROSITE; PS50837; NACHT; 1.
SQ   SEQUENCE   892 AA;  101733 MW;  AE703D8DF341C2AC CRC64;
```

```
Query Match             18.2%;  Score 1371.5;  DB 4;  Length 892;
Best Local Similarity   35.2%;  Pred. No. 2.6e-94;
      Matches  356;  Conservative  135;  Mismatches  349;  Indels  171;  Gaps  22;

Qy    8 RLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETP-AQPEKTSGMEVASYLVAQYGEORA 66
        :|::|:|:  |::|:  |:  |:||  |:::  |::::::|  ||::   :||::||:::|
Db   12 RLSTYLEELEAVELKKFKLYLGTA--TELGEGKIPWGSMEKAGPLEWAQLLITHFGPEEA 69
        ...
```

(sequence alignment continues)

```
Db   245 DYLFYINCREMNQSATBCSMQDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDELKPSFH 304
Qy
Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLBEQARWEVLG 480
Db
Db   305 DPQGPWCLCWBEKRPTELLLNSLIRKKLLPELSLLIITTRPTALEKLIHRLLEHPRVEILG 364
Qy
Qy   481 FSESSRKEYFYRYRYFTDERQAIRAFRLIVKSNKELWALCLVPWVSMLACTCLMQQMRKEKL 540
Db
Db   365 FSEAERKEYFYKYFHNAEQGQVFNYVRDNEPLFTMCFVPLVCWVVCTCLQQQLBGGGLL 424
Qy
Qy   541 TLTSKTTTLCLIHYLAQALQAQPLGPQL------RDLCSLAABGIWQKTLFSPDDLRKH 594
Db
Db   541 VTRLLTEYAPSERSFLALTSRFLFGLLNEBTRSHEBKSLCWKVSPHIKDLLQWIQSKAQ 600
Qy
Qy   425 RQTSRTTTAYYMLYLSLLSMQPKPGAPRLQPPPNQRGLCSLAAADGIWNQKILFEBQDLRKH 484
Db
Db   485 GLDGEDVSAFLNMNIFQKDINCERYYSFIHLSFQEFFAAMYYILDEGEGAGPDQ----D 540
Qy
Qy   595 GLDGAIISTFLKMGILQEH-PIPLSYSFIHLCFQBFFAAMSYVL-EDEKGRGKHSNCIID 652
Db
Qy   653 LEKTLEAYGI-HGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ- 708
Db
Db   601 SDGSTLQQGSLEFFSCLYEIQEBEBFIQQALSHFQVIVSNIASKMEHMVSSFCLKRCRSA 660
Qy
Qy   709 ---LLIQHSLESLHCLYETRNKIFLITQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHV 764
Db
Db   661 QVLHIYGATIYSADGEDRARCSAGAHTLLVOLRPERTVLLDAYSEHLAAALCTNPNLIELS 720
Qy
Qy   765 KKLQLI----------EGRQHRSTWSPTMVLLFR--WVPVTDAYWQILFSVLKVTRNLKELD 814
Db
Db   721 LYRNALGSRGVKLICQGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 780
Qy
Qy   815 LSGNSLSHSAVKSLJCKTLRRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTLITELDLSF 874
Db
Db   781 NGVGFPPGMMLLCEGLRHPQCRLIQMTQLRKCQLESGACQEMASVLGTNPHLVEDLITGNAL 840
Qy
Qy   875 NVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNL 934
Db
Db   841 EDLGLRLLCQGLRHPVCRLRTLWWLFGMD---LNKMTHSRLAALRVTKPYL 888
Qy
Qy   935 DDVGVRLLCEGLRHPACKLIRL----GLDQTTLSDEMRQELRALEQEKPQL 981
Db
```

```
Query Match  16.7%;  Score 1254.5;  DB 11;  Length 1033;
Best Local Similarity  32.6%;  Pred. No. 2.2e-85;  SMW; 118274 MW;
Matches 331;  Conservative 164;  Mismatches 334;  Indels 187;  Gaps 23;

Qy     8 RLACYLEFLKKEELLKEFQLLLANKAHSRSSSGETPAAVSYILVAQYGEQRAW 67
Db     7 KLAQVLEDLEDVDLKKKFKWMHLEDYPPEKGCIPVPRGQMEKADHLDLATLMIDFNGEBKAW 66
Qy    68 DLALHTWBQMGLRSLCAQAQGEGAGHSPSFPYSPSEPHLGSPSQPIPSTAVLMPWIHELPAG 127
Db    67 AMAVWIFAAINRRDLWEKAKK-----------------DQP--------EWNDT 95
Qy   128 CTQGSERRVLRQLPDTSGRWREISASLLYQALPSSPDHBSPSQBSSPNAPTSTAVLGSWG 187
Db    96 CTSHSS-MVCQE--DSLEEEW----------------MGLLG 118
Qy   188 SPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIREBEREKSEKGRPPWAAVVGTPPQAHTS 247
Db   119 YLSRSICKKKK-----------DYCKMYRHVRSRFYSIKDRN----ARLG--- 155
Qy   248 LQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSWPDVEEN 307
Db   156 ------ESVDLNSRYTQLQLV-KSHPSKQE------ 178
Qy   308 RGHLI-----EIRD---------LFGPGLDTQEP-RIVTLQGAAGIGKSTLARQVKEA 350
Db   179 ------LGPQLRDLCSLAAABGIWQKI 238
Qy   351 WGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILD 410
Db   239 WALGKLFKDKPDYLFPIHCREVSLRTPRSLADLIVSCWPDENPFPVCKILRKPSRILFLMD 298
Qy   411 GVDEPGWVLQBPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSL 470
Db   299 GFDELQGAFDEHIGEVCTDWQKAVRGDILLSSLIRKKLLPKASLLITTTRPVALEKLQHLL 358
Qy   471 EQARWEVLGFSESSRKEYFYRYRYFTDERQAIRAFRLIVKSNKELWALCLVPWVSMLACTCL 530
Db   359 DHPRHVEILGFSEAKRKEYFFKYFSNELQAREAFRLIQENEVLFTMCFIPLVCWIVCTGL 418
Qy   531 MQQMRKEKLTLTSKTTTLCLIHYLAQALQOAQP------LGPQLRDLCSLAAABGIWQKKT 584
Db   419 KQQMETGKSLAQTSKTTTAVVFFLSSLLQSRGGIBEHLFSDYLLQQLCSLAAADGIWNQKI 478
Qy   585 LFSPDDLRKHGILDGAIISTFLKMGILQ-EHPIPLSYSFIHLCFQBFFAAMSYVLEDE-- 640
Db   479 LFEBCDLRKHGILQKTDVSAFLRMNVFQKEVDCERYSFSHMTFQEFFAAMYYLLEBEAEG 538
Qy   641 ----KGRGKHSNCI-IDLEKTLEAYG-IHGLFGASTTRFLLGLLSDEGEREMENIFHCRL 694
Db   539 ETVRKGPGGCSDLLNRDVKVLLENYGKFBKGYLIFVVRFLFGLLVNQERTSYLEKKLSCKI 598
Qy   695 SQ--GRNLMQWV----PSLQLLLQPHSLESLHCLYETRNKIFLITQVMAHFEEMGMCVETD 748
Db   599 SQQVRLELLKWIBVAKAKLQWQPSQLELFYCLYEMQEEDFVQSAMDHFPKIEINLSTR 658
Qy   749 MELLVCTFCIKFSRHVKKLQI-------IEGRQHRSTWSPT------MVLLFRWV 790
Db   659 MDHVVSSFCIKNCHRVKTLSLGFFHNSPKEEEEBRRGGRPLDQVQCVFPDTHVACSSRLV 718
Qy   791 P---VTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCG 848
Db   719 NCCLISSFCRGLFSSLSTNRSLTELDLSDNTLGDPGMRVLLCEALQHPGCNIQRLWLGRCG 778
Qy   849 LTAEDCKDLAFGLRANQTLITELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTS 908
```

DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00368; LRR_RI; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS00038; HLH_1; 1.
DR   PROSITE; PS50503; LRR_RI; 3.
DR   PROSITE; PS50837; NACHT; 1.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR001092; HLH.basic.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00560; LRR; 2.

```
Db   779 LSHQCCFDISSVLSASPSLKELDLQQNNLDDVGVRLLCEGIRHPACKLIRLGLDQTTLS 838
             |||   :::||||||||||||:| ||  |  ||||| | | | | ||||||::| | : ||
Qy   909 DCCQDLASVLSASPSLKELDLSDNALGDFGIRLLCVGLKHLLCNLQKLMLVSCCLTS 964
Db   839 ACCQDLALVLSSNHSLTRLYIGENALGDSGVQVLCERMDPQQCNLQKLGLVNSGLT 894

Qy   256 EPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPH----- PRSQDPLVKRSWPDYVBENRG 309
             |||:| ::|   ||||| : | ::: ||
Db     9 GLMWYLEELNKKEFVKFEFLKQEVLQLGLKQVSWTEVKKASRQDLASLLLKIHYBEKPAW 68

Qy   203 GTQWPLDETSGIYTEIRE-RERERKSEKG--RPFWAAVVGTPPQAHTSLQPHHH-----PW 255
             |  ||  :::   : |  |    :|| |    : :  | : |  ||: |
Db   114 FTISIQNFLKEKFTBDDYDCFENLFQSKGTESKPQVVFLSGGAGVGKTLMLKRLMLAWTE 173

Qy   310 HLIEIR----------- DLFGPGLDTQEPRIVILQGAAGIGKSTLAROVKEAWGR 353
             ::|:|              :|| | : |:|   | :|| | :: | | ::||
Db    69 DMTFR------------PFQKINRKDLIERAKREIDGGCPKLYRAHMKTKMTH--DSSRA 113

Qy   354 GQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVD 413
             |  : |  |   | :|  :|| | :: | |   :  |  | | |  ||||||:||:
Db   174 SPVFLHKFSYIFYFPCCREVKQLKTASLAELISREWPGFSAPIEEILSKPEKLLPIIDSLE 233

Qy   414 EPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLBQA 473
             | |: |  |||:| |        | | : | | :: || :|   ||  |  ||  |:
Db   234 GMECDLFKWESELCDNCTEKQPVNVLLSSLIRRKMLPESSLLISATPESFSPEKMENRIEYT 293

Qy   474 RWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLLVPWVSWLACTCLMQQ 533
             |  || | ||| |   |  :  ::| |:| |    || | |: ||  ||  ||  |  :
Db   294 HVKIIKGLKERNIKMSFHRLFQDRNRAHEAFSLVRENEQLFTVCQVPVLCWMVATCLKEE 353

Qy   534 M-KRKEKLTLTSKTTTLCLHYLAQALQAOPLGP-------QLRDLCSLAAEGIWQKKTLF 586
             | |::    :| |||: : :|  |  |:  :: |        |||  | : | | |:: |
Db   354 IEKGRDPVSICRCTTSLYTTHIFNLFIPQNAHSPSKSQODLQGLCSLAAEGNWTDTFVF 413

Qy   587 SPDDLRKHGLDGAIISTFLKMGIL--QEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGK 645
             || | : |  | |:|: || ||:|    || |||  || |: :: | |:| |:  | |
Db   414 GBEALRNGIMDSDIPILLDIGMLINIRESEKSYIFLHPSVQEVLCAAIFYLL----KSHVD 470

Qy   646 H-SNCIIDLEKTLEAV----GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GR 698
             | |    :| | ||      |: | : | |  |||| |::   ||||: | ||  |   ||
Db   471 HPSQEVKSIEKLMFAFLKKVKVQWIFFGS---FIFGLLHESEQKKLEAPFGHQLSQEIKR 527

Qy   699 NLMQWVPSL-----QLLLQPHSLSESLHCLYETRNXTFLTQVMAIFEEMGMCVEFDMELLVC 754
```

RESULT 5

Q8BU40

Q8BU40;

ID Q8BU40; PRELIMINARY; PRT; 982 AA.
AC Q8BU40;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to PAN2 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK087843; BAC40024.1; -.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFEA2053585 CRC64;

Query Match 11.9%; Score 899.5; DB 11; Length 982;
Best Local Similarity 26.8%; Pred. No. 1.2e-58;
Matches 269; Conservative 176; Mismatches 366; Indels 191; Gaps 26;

```
Qy   528 QLYQCLETISGNEELQEQIDGMKLFYCLFEMDDDTFLVEAMNCMEQINFVAKDYSDVIVA 587
             : :::   |:::|  ||:|    | |  |:  |  |::        ::| |   : :|
Db   793 TDAVWQILFSVLKVTR------------------------NLKELDLSGNSLSHSAV 825
Qy   755 TFCIKFSRHVKKLL-------QLIEGRQHRS------TWSPTMVVL-------PRWVPV 792
             ::| : : : |              | :| :        || :| :         |  |
Db   588 AHCLKHCFLKKLSFSTQNVLSGAQEHSTMERLLTAWNICSVFIISKDIQELRMKDTNL 647
Qy   826 KSLCKTLRRPRCLLETTLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHL 885
             |||  |||  |   |:|| :   || :|  ||:  | ||| |: |||:|:|| ||   |
Db   648 SESAFSVLYNNLKYHNYTLNVLVANNVFFVCEKYLFFELIQNCNLQHLLNLSLTILSHSDV 707
Qy   708 KLLCDVLSQAECNIEELVVAACSISSDDCKVFASVLISNKTLKHLNLASNTL-DKGIASL 766
             |:|  |||: |   ||| || :|| |  |  | :|| |  || ||:|:| :|
Db   826 CKSLCHPDCILEHLVLANCSLNEKCWDYLSEVLRNKTLSHLDISSNDLKDEGLKVLCRA 826
Qy   949 PACKLIRLGLDQTTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTS 1008
             |||||||||||| ||   ::| |  ||:|| | : :||  ||  ||| || || ::  |
Db   767 PNCHLEDLGLBACELTGACCEDLASTFTQCKTLWAMNLLK------NALDYNGLVVLCE 939
Qy  1009 SLKRQRLGSERAASHVAQANLKLLDV-SKIFPIAEIAESSP 1049
             ||: |:||   ||  :|| || ||   :|| |:| ||  |
Db   940 ALKQQ----TCATYVLGLQITDFDTETQAFLVAE--QEKNP 974
```

RESULT 6

Q8C6J9

Q8C6J9;

ID Q8C6J9; PRELIMINARY; PRT; 863 AA.
AC Q8C6J9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical RNI-like structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054426; BAC35775.1; -.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 99955 MW; D3A46C782DA544CF CRC64;

Query Match 11.5%; Score 868; DB 11; Length 863;
Best Local Similarity 27.7%; Pred. No. 2.4e-56;
Matches 236; Conservative 154; Mismatches 343; Indels 120; Gaps 17;

```
Qy   214 IYYTEIRERERE------------KSEKGRPPWAA------VVG---TPPQAHTSLQPHHHPWEP 257
Db    41 ISWTEVKEASREDLAILLVKHCDGNQAWDTTFRVFQMIGRNVITNRATGEIAAHSTIYRA 100
Qy   258 SVRESLCSTWWPWKNEDFNQKFTQLLLLQRPHPRSQODPLVKRSWPDYVEENRGHHIEIRDL 317
Db   101 HLXEKL------THDCSRKP-----NISIQNFFQDEYDHL---ENL 132
Qy   318 FGPGLDTQEPRIVILQGAAGIGKSTLAROVKEAWGRGQOLYGDRFQHVFYFSCRELAQSKV 377
Db   133 LVPNGTENNPKMVVLQGVAGIGKTILLRNIMIVWSEGLVPQNKFSYIFYFPCHDVKQLQT 192
```

```
RESULT 7
Q8IXT0
ID   Q8IXT0        PRELIMINARY;      PRT;   846 AA.
AC   Q8IXT0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to NALP2 protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Strausberg R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC039269; AAH39269.1; -.
SQ   SEQUENCE   846 AA;   96368 MW;   46BB824550E39F7 CRC64;

Query Match     11.2%;   Score 843;   DB 4;   Length 846;
Best Local Similarity   30.0%;   Pred. No. 1.8e-54;
Matches 255;   Conservative 154;   Mismatches 328;   Indels 114;   Gaps 28;
```

```
Qy   378  VSTLAELIGKDGTATPAPIRQILSRPERLLFILLDGVDEPRGWVLQEPSSELGCLLHWSQPQPQPAD  437
          ||:||  ::     : .:::.:||||:|:||:||..||||..:||:|.| .::.:
Db   193  ASLADLISREWPSSAPMEILLSQPEKLLFIIDSLEGMEWNVTQQDSQLCYNCMEKQPVN  252

Qy   438  ALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFDE  497
          .||.|||.|.||| ::|| ||  : . |.: : .  :|:  : : .||.|.|. :
Db   253  VLLSSLLRKILPESSLLISTSCETFKDLKDWIEYTNVRIITGFKENNINMCFHSLFQDR  312

Qy   498  RQAIRAFRLVKSNKELWALCLIVPWVSWLACTCLMQQMKR-KEKLTLTSKTTTTLCLHYLA  556
          . |:.||.|:.:|:||| .| . .|::.: |||| ::.  :|.|  .|  ||  |.|| :
Db   313  NIAQEAFSLIRENEQLFTVCQAPVVCYMVATCLKNEIESGKDPVSICRRTTSLYTTHILN  372

Qy   557  QALQAQPLGPQ------LRDLCSLAAAEGIWQKKTLFSPDDLRKHIGLDGAIISTELKMGII  610
          . :.| :  :       |.:||.||...| |.:  :|.||.::.:.|.. .|||..| ||:
Db   373  LFIPHNAQNPSNNSEDLLDNLCFLAVEGNWMTDISVFNERALRRNGIMDSDIPTLLDIGIL  432

Qy   611  -QEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIID------LEKTLEAYGIH  663
           :. .: .|.:.:.||||:|.|||:|.|: ... . . :.||:      ||:|:.:..:|
Db   433  EQSRESENSYIFLHPSVQEFCAAMYYLLHSEMDHSCQGVVFIETFLFTELNKIKKQWVFL  492

Qy   664  GLFGASTTRFLIGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQLLLQP----HSL  716
          | ..  : :||:||| .||: : ...    ||.  | : :   . :.|:. :    |||
Db   493  GC--------FFPGLLHETEQEKLEAFFGYHLSKELRRQLFLWLELLLDTLHPDVKKINTM  545

Qy   717  ESLHCLYETRNKTFLITQVMAHFEEWGMCVETDMELLLVCTPCIKF-----SRHVKK  766
          .: .|  : :|:  .  .|  .|::   .:. |::..  . |    :     ||.||
Db   546  KFFYCLFEMEBEVFVQSAMNCREQIDVVVRGYSDPIVAAVCLSHGSALTDFSISAQONVLN  605

Qy   767  LQLIEGRQHRSTWSPTMVVLFRWPV---------TDAYWQILFSVLK---------  805
          .:|:.|:.   .:.  . .:|: .:         ||.:  :||||:|
Db   606  EELGQRGKLILILWHQICSVFLRNKDIKTLRIEDTIFNEPVFKIFYSYLKNSSCILKTLVA  665

Qy   806  --------VTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGL  849
                  :|||||:|||| ||||:||.|.:||  |||| ||||||||||||
Db   666  YNVSFLCDKRLFLELIQSYNLEELYLRGTFLSHSDVEMLCDILNQAECNIRILDLANCSL  725

Qy   850  TAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSD  909
          .   ||||| :||.|.:|.|::..:|.|.|.|  .|..:|  ||.:|::|.| . | |.|
Db   726  CEHSWDYLSDVLRQNKSLRYLINISYNNLKDEGLKALCRALTLPNSALHSLSLBRACQLTGA  785

Qy   910  CCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLLIRLGLLDQTTLSDEMRQ  969
          ||:|||| : :.| ||::|| :||||.||:.:| .:|| |:||::.| |  :. .|||:.:
Db   786  CCKDLASTFRYKCLRRINLAKNSLGFSGLFVLCKAMKDQTCTLYELKLRMADF-DSDSQ  844

Qy   970  ELRALEQERKFQLL  982
          ::.  .|||.|.:|
Db   845  EFLLSEMERNKIL  857
```

```
Qy   148  WREISASLLYQALPSSPDHESPSQESPNAPTSTAVLGSWGGSPPPQOPSLAPREQEAPGTQWP  207
          | .:|||::.:     .  . .: :   :.  | :.|    :.:   .:.
Db    67  WVEMASLQVFEKM-----HRMDLSERAKDEVREAALKSFNKRKPLSLGITRKERP---P  117

Qy   208  LDETSGIYYTEIRERERKSEKGRPPWAAVGTPPQAHTSLQPH-----HHPWEPSVRES  262
          ||:|    : ..|.|:    |.      |.. :.. .|| |       :.||.|.|||
Db   118  LD-----VDEMLERRTEAQAFTETKGNVICLGKEVFKGGKKPDKDNRCRYIILKTKFRE-  170

Qy   263  LCSTWPWKNEDFN--QKFTQLLILQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFG  319
          :.|.|| ::.::.  ::. :|   ..: ..  | .|  |||| ||.|||:|||||||||
Db   171  MWKSWPGDSKEVQVMAERYKMLIPFSNP------RVLPG  203

Qy   320  PGLDTQEPRIVILQGRAAGIGKSTLARQVKEAWGRGQLYGDRFQHVYFSCRELAQSKVVS  379
          | .:|:||||::|.| | |:|||||| |:   |  .:|  .:|::|:||||||::. .||
Db   204  PFSYT-----VVLYGPAGLGKTTLAQKLMLDWAEDNLI-HKFKYAFYLSCRELSRLGPCS  257

Qy   380  LAELIGKDGTATPAPIRQILSRPERLLFILLDGVDE-----PGWVLQEPSSELGCLLHWSQPQP  435
          :||: :.| :. .  .|:.|:. |:.||::.||.||      |.  :.::.||| .::
Db   258  FAELVFRDWPELQDDIPHILAQOARKLLFVIDGFDELGAAPGALIB---DICGDWEKKKP  313

Qy   436  ADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFT  495
          : .:||||| .::.| |.| :|.|  .||:|  .| : |:|:: : : .:: :| |: |
Db   314  VPVLLGSLLNRVMLPKAALLVTTRPPALRDLRILAEEPIYIRVEGFLEEDRRAYFLRHFG  373

Qy   496  DERQAIRAFRLVKSNKELWALCLIVPWVSWLACTCLMQQMKRKEKLTLTSKTTTTLCLHYL  555
          ||.||:||| :.|||.|| .| : .|::.:. |||:. :...:|  |. .|: ||.||||
Db   374  DEDQAMRAFELMRSNAALFQLGSAPAVCWIVCTLKLQMEKGEDPVPTCLRTGLFLRFL  433

Qy   556  AQAL-OAQPLGPOLRDLCSLAAEGIWQKKTLFSPDDLRKHIGLDGAIISTELKMGIIQEHP  614
          .. | .||:.  |||:||||||.| |.: .:|.::||::.:. .. ::.. : :.::.:
Db   434  CSRFPQGAQLRGALRTLSLLAAQGLWAQTSVLHREDLERLGVQESDLRLFLDGDILRQDR  493

Qy   615  IPLS-YSFIHLCFQEFFAAMSYVL--EDEKGRGKISNCIIDLEKTLEAYGIHGLFGASTT  671
           .:. |||||| |.:| .|:.|.|  |:|:.  .     |:|:|.|:.:| :|||.|.|
Db   494  VSKGCYSFIHLSFQQFLTALFYTLEKEEEDRDGHTWDIGDVQKLLS--GVERLRNPDLI  551

Qy   672  R---FLLGILSDEGEREMENIFHCRLSQGRNLMQWVPSL-QLLLQ------PHSL  716
          :   ||| ||| .||: :.:.    |::| : :.   .| :::|.      |||
Db   552  QAGYSFGLANEKRAKELERATFGCRMS------PDIKQELLRCDISCKGGHSTVTDL  602

Qy   717  -ESLHCLYETRNKTFLTQVMAHFEEWGMCVETDMELLLVCTFCIKFSRHVKK--LQLI-EG  772
           ::| ||  : :|:  . .|  .|::   .:. |:...  .|    ||:|:   |||: .
Db   603  QELLGCLYESQBEELVKEVMAQFKEISLHLNA-VDVVPSSFCVSHCRNLQKMSLQVIKEN  661

Qy   773  RQHRSTWSPTMVVLFR------WPVTDAYWQILFSVLKYTRNLKELDLSGNSLSHSAVK  826
           :.:|.:.  . .:|:      || ||.:  :|.|:  .::|::: .:  :.|| |::|
Db   662  LPENVTASESDAEVERSQDDQHMLP-----PWTDLCSIFGSNKDLWGLAINDSFLSASLVR  717

Qy   827  SICKTLRRPRCLLETLRLAGCGLTRAEDC-KDLAFGLRANQTLTELDISFNVLTDAGAKHL  885
          .|.: .:.   .:| ..:.  .| :||:  :.:|: |.:.::.. |: .:|.:. :|| .:
Db   718  ILCEQIASDTCHLQ--RVVFKNISPADAHRNICLALRGHKTVTYLTLQGNDQDDM-FPAL  774

Qy   886  CQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEG  945
          |  ||.|.| | ||.||||. :.   .||: .. .    . .:|.: : .  |. :: :
Db   775  CEVLRHPECNLRYLGLVSCSATTQQWADLSIALEVNQSLTCVNLSDNELLDBGAKLIYTT  834

Qy   946  LRHPACKLIRL  956
          ||||.|. :||
Db   835  LRHPRCCFLQRL  845
```

```
RESULT 8
Q95L27
ID   Q95L27        PRELIMINARY;      PRT;   713 AA.
AC   Q95L27;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 80.3 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
```

Left column

```
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Hashimoto K., Obada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA   Terao K., Sugano S.;
RT   "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT   libraries.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB071044; BAB64437.1; -.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007091; LRR_RI; -.
DR   PROSITE; PS50503; LRR_RI; 2.
DR   PROSITE; PS50837; NACHT; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   713 AA;  80258 MW;  6F214C9B773F54DC CRC64;
```

Query Match

Best Local Similarity   10.8%;

Matches 228; Conservative 114; Mismatches 248; Indels 135; Gaps 18;

Score 815.5; DB 6; Pred. No. 1.6e-52; Length 713;

```
Qy    383 LIGKDGTATPAPIRQILSRPERLLFILDGVDE----PGWVLQPPSSELCLHWSQPQPADA 438
Db      1 MVFRDWPELQDDIPHILAQAQKILFVIDGFDEGAPPGALIQ----DICGDWEQQKPVPV   56

Qy    439 LLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFTDER  498
Db     57 LLGSLLKRKMLPKATLLVTTRPRALRDLRFLAEQPIYIRVEBGFLEEDRAYFLRHFGEDD 116

Qy    499 QAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTTLCLHYL-AQ  557
Db    117 QAMRAFELMRSNNAALFQLSGSAPAVCWIVCTTLKLQOMEKGEDPAPTCLTSTGLFLRFLCSQ 176

Qy    558 ALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPL 617
Db    177 FPGQAQLRGALRALSLLAAAQSIWAQMSVLHGEDLESAGVQESDLRLFLDGDILRQDGVAK  236

Qy    618 S-YSFIHLCFQEFFAAMSYVL--EDEKGRGKHSNCIIDLEKTL---EAYGIHGLFGASTT  671
Db    237 GCYSPIHLSFQOPLTALPYALEKEEEEDRDGHAWDIGDVQKLLSREBRLKNPDLIQAG--  294

Qy    672 RFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSL-QLLLQ------PHSLES       718
Db    295 RFLFGLANEKRVKELEATFGWRMS---------PEIKQELLRCDVSRKNGHFTAADLREL  345

Qy    719 LHCLYETRNKTFLTQVMAHFEEKMGMCVETDMELLVCTFCIKFSRHV------KKLQL-  769
Db    346 LCCLYESQEDELVKEWMAQFKEISLHLNA-VDIAPSSPCFKHQCNIFPADACRNLCLALR  404

Qy    770 -------IEGRQHRSTWSPTMWVL----FRWV------PVTDAYWQILFSVLKVTRN    809
Db    405 GHKTVTHLTLQGTDQKDMLPALCEVLRHPBCNLRYLGLVSCSAITTQQWADLSLALBANRS 464

Qy    810 LKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGGLTAEDCKDLAFGLRANQTLTE  869
Db    465 LMCVNLSDNELLDDEGAKLLYTTLRHPKCFLQRLSLENCHLTEANCKDLAAVLVVSRELTH 524

Qy    870 LDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLLTSDCCQDLLASVLSASPSLKELDL 929
Db    525 LCLAKNSLKDTGVFLCEGLSYPECKLQALVLWNCDITSDGCCSLAKLLQEKSSLSCLDL   584

Qy    930 QQNNLDDVGVRLLCEGLRHPACKL-----                               953
Db    585 GLNHIGVTGVGKVLCEALSKPLCNLRCLWLWGCSIPPFSCEDLCSALSCNQSLSTLDLGQN 644

Qy    954 -----IRLGLDQTTLSDEMRQELRALEQEKPQLLIFS------                 985
Db    645 PLGSSGVRMLFKTLTRPGTLQTLRIKIDD--FNDELHKLLEEIBENNPQLIIDTEKDPW  702

Qy    986 RRKPS 990
          ::||||
Db    703 KKRPS 707
```

Right column

```
RESULT 9
Q99MW0

ID   Q99MW0      PRELIMINARY;      PRT;   748 AA.
AC   Q99MW0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ribonuclease/angiogenin inhibitor 2.
GN   RNH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21175748; PubMed=11279525;
RA   Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT   "An abundance of X-linked genes expressed in spermatogonia.";
RL   Nat. Genet. 27:422-426(2001).
DR   EMBL; AF285581; AAK31960.1; -.
DR   HSSP; P10775; 2BNH.
DR   MGD; MGI:1890518; Rnh2.
DR   InterPro; IPR007091; LRR_RNinh.
DR   PROSITE; PS50503; LRR_RI; 1.
SQ   SEQUENCE   748 AA;  84946 MW;  674A385E011DE8EC CRC64;
```

Query Match

Best Local Similarity   9.2%;

Matches 197; Conservative 120; Mismatches 241; Indels 193; Gaps 14;

Score 689.5; DB 11; Pred. No. 5.6e-43; Length 748;

```
Qy    419 LQEPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEV  478
Db      5 MSBRESELCDTCTEKQPLRILLSSLLRRKMLPKSSFLISATPETFEKMEGRVBECTNVKIV  64

Qy    479 LGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQM-KRK  537
Db     65 TGFNESNIKMYFRSLFQDKTKTQEIFSLVKRNQOLFTVCQVPVLCWMVAATCLKKEIEKGR 124

Qy    538 EKLTLTSKTTTTLCLH----YLAQLQ--AQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDL  591
Db    125 DIVSVCRRTSLYTTHIRNLFIPQSAQYPSKESQOAQLQSLCLAAAGMWTDTFVFGEEAL   184

Qy    592 RKHGLDGAIISTFLKMGILQE-HPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCI  650
Db    185 RRNGIMDSDIPTLLDVRILEKSKKSSEKSYIFLHPSIQEVCAAIFYLL---KSHMDHPSQD 241

Qy    651 IDLEKTLEA---------YGIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNL 700
Db    242 V---KSIEALIFTFLKKVVQWIFFGS---FIFGLLHSBEQKLEAFGHQLSQBIKRQL     295

Qy    701 MQWVPSL----QLLLQPHSLESLLHCLYETRNKTFLTQVMAHFEEKMGMCVETDMELLVCTF 756
Db    296 YQCLETISGNEELQEQVDGMKLPYCLFEMDDEAFLAQANNCMEQINFVARDYSDVIVAAH  355

Qy    757 CIKFSRHVKKLQ-----LIEGRQHRST-----W                           779
Db    356 CLQHCSTLKKLSLSTQNVLSEGQEHSYTEKLLMCWHHMCSVLISSKDIYILQVXNTNLNE 415

Qy    780 -----                                                       780
Db    416 TASLVLYSHLMYPSCTLKALVVNNVTFLCDNRLPFELIQNQCLQHLDLNLTFLSHGDVKL 475

Qy    780 -----                                                       780
Db    476 LCDVLSQEECNIEKLMVAACNLSPDDCKVFASVLISSKQWLKHLNLSSNNLDKGITSSLSKA 535

Qy    791 -----PVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRR           834
Db    536 LCHPDCVLXNIVLVNCSLSEQCWDYLSEVLRRNKTLNHLDISSNDLKDEGLKVLCRALSL 595

Qy    835 PRCLLETLRLAGGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSC 894
```

```
Db   596 PDSVLKSLSVRYCLITTSGCQDLAEVLRKQNLRNLQVSNNKIEDAGVKLLCDAIKHPNC 655
         ||:| :||: ||  || |||:|| |||||  |::| |:|:||||||:|||:| ||:||:|
Qy   895 KLQRLQLVSCGLITSDCCQDLASVL.SASPSLKEILDQQNNLDDVGVRLLCEGLRHPACKLI 954

Db   656 HLENIGLEACALTGACCEDLASAFTHCKTLWGINLQENALDHSGLIVLFEALKQQOCTLH 715
Qy   955 RLGLDQTTLSDEMRQELRALEQEKPQLLIFS 985

Db   716 VLGLRITDFDKETQELIMAEEEKNPHLSTLS 746

Db   514 GVSKXYSIQNMLSREARLKNPDLSGLL----------PFVFGLLNETRIQELKTTFGCQISTEVKR 568
Qy   696 -----QGRNLMQWVPSLQLILQPHSLESLHCLYETRNKTFLTQVMAHFEEMGMCVETDM 749

Db   569 KFLECESGEN--------KPLLILMNMQEILSCLYESQSEGFVKRAMVLFEDISLHLKTST 621
Qy   750 ELLVCTFCIKFSRHVRKLQL-IE------------GRQHRSTWSPTMVVLFRWVPVTD 794

Db   622 DLIHASFCLKNSQNLQTMSLKVEKAVFPENVAALESTAKHQRSPDEQRML----- 671
Qy   795 AYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDC 854

Db   672 TFWTDFCDTFNSNKKLVFLDIHESFLNSSALEILCEKLPSASCCLQKVVLK--NISPDDA 729
Qy   855 -KDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLITSDCCQD 913

Db   730 YEKULCLIFNGYKTISHLILQGGNL-DSMHHSLCEVLKNPACNLKPFLSLGSCSTAAQKWDD 788
Qy   914 LASVLSASPSLKEILDLQQNNLDDVGVRLLCEGLRHP 949

Db   789 FFPVLKVNQSLIFLDLTDNSLLDKSAKLLCNIWKEP 824
```

```
Query Match             9.0%;  Score 679.5;  DB 11;  Length 825;
Best Local Similarity  26.0%;  Pred. No. 3.7e-42;
Matches  212;  Conservative 150;  Mismatches 313;  Indels 141;  Gaps 24;

Qy   208 LDETSGIYYTEIRERERERKSEKGRPPWAAVVGTPPQQAHTSLQPHHIHPWEPSVRESLCSTW 267
Db    76 LEEINRVDLAELLVKQIEB-------------AVLKVPEEKVSS----KPREPSG----TLTF 117

Qy   268 PWKNEDFNQKFTQLLLLQRPHPHRPSQDPLVKR--------SWPD------YVEENRGHLIE 313
Db   118 PW------NFVQ--GAKRPEDKQKEKWKTRYTAKWKQNFWPKCNKEIYVVIESYKTLLA 168

Qy   314 IRDLFGPGLDTQGPRIVILQLQGGAGIGKSTLARQVKEANGRGQLYGDRFQHVFYFSCRELA 373
Db   169 ---LCNPKIETPFAHAIVLHGPPGSGKTTMAKQLMLEWSESK-QAQIFSCAFYISCREVN 224

Qy   374 QSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQP 433
Db   225 NTKPCTFAHLLSMDNPSWRDCVIRDLLIGKEFLFVVDGFDELFPAGALIRDLCGDWNTV 284

Qy   434 QPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYRY 493
Db   285 KPVEVLLGSLLKRKMAPHATLLVTTRTQSLHQIFVMMDQPLLVETLGFLEQEKQEYFQXY 344

Qy   494 FTDER-------QAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRBKLTLTS 544
Db   345 FEDEGEGEEDKGEGKALRALKEVRCNADLYQMASLPTACGIFCLCLELRMKKGEDLSLTC 404

Qy   545 KTTTTLCLHYLAQALQAQPLGPQLRD-------LCSLAAEGIWQKKTLFSPDDLRKHGL 596
Db   405 QTYTSMFLNFLCEVFSSETCEDHLNEEFQILFKICILAANSLLEQVPLLCEEDF----- 459

Qy   597 DGAIISTFLKMGILLQEHPIPLSY-----SFHILCFQEFFAAMSYVLE---DBK 641
Db   460 ------LTLKLNLNLHPMVCRHILFKDSSSTHCLSFICLGIQOLLAAIIFVQELGQESK 513

Qy   642 GRKGHS-NCIIDLEKTLBAYGIHGLFGFASTTRFLLGLLSDEGEREMENIFHCRLS----- 695
```

```
Query Match             9.0%;  Score 676;  DB 11;  Length 657;
Best Local Similarity  27.1%;  Pred. No. 4.8e-42;
Matches  194;  Conservative 95;  Mismatches 220;  Indels 206;  Gaps 11;

Qy   314 IRDLFGPGLDTQE-PRIVILQGGAGIGKSTLARQVKEANGRGQLYGDRFQHVFYFSCREL 372
Db    78 LQHIFDEDIQTSEAPQTVVLQGAAGIGKTTLLKKAVLEWADGNLY-QQFTHVFYLNGKEI 136

Qy   373 AQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQ 432
Db   137 SQVXEKSFAQLISKHWPSSEGPIEQ------------GQVEKGQRVEE------- 172

Qy   433 PQPADALGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYR 492
Db   173 -------------------------------------- 172

Qy   493 YFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKREKLTLITSKTTTTLCL 552
Db   173 --------------------------------TCKTTSTALFT 182

Qy   553 HYLAQALQOQPLG------POLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFL 605
```

RESULT 12

```
ID   Q8CCN1;          PRELIMINARY;     PRT;   673 AA.
AC   Q8CCN1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   PAAD and NACHT containing protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium.
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs."
RL   Nature 420:563-573(2002).
DR   EMBL; AK032446; BAC27872.1; -.
SQ   SEQUENCE   673 AA;  76367 MW;  EE773C592BEC7054 CRC64;
```

Query Match   7.6%;   Score 573;   DB 11;   Length 673;
Best Local Similarity   33.9%;   Pred. No. 2.9e-34;
Matches 152;   Conservative 81;   Mismatches 154;   Indels 62;   Gaps 14;

```
Db  183 YYICSLFPRIPVGCVTLPNETLLRSLCKAAVEGIWTMKHVLYQQNLRKHELTREDILLFL 242
Qy  606 KMGILQEH-PIPLSYSFIHLCFQEFFAAMSYVL-EDEKGRGKHISNCIIDLEKTLEAYGIH 663
Db  243 DAKVLQQDTEYENCYMFFHLHVQEFFAALPYLLRENLEEQDYPSEPFENLYLLLESNHIH 302
Qy  664 GLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ---LLLQPHSLE 717
Db  303 DPHLEQMKCFLFGLLNKDRVRQLEETFNLTISMEVREELLACLEGLEKDDSSLSQLRFQD 362
Qy  718 SLHCLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRS 777
Db  363 LLHCIYETQDQEFFITQALMYFQKIIVRVDBEPQLRIYSFCLKHCHTLKTMRLTARADLKN 422
Qy  778 TWSPTMVVLFRWVPVTDAYWQILFSVLLKVTRNLKELDL--------------- 815
Db  423 MLDTAEMCLEGAAVQVIHVWQDLFSVLHTNESLIEMDLYESRLDESLMKILNEELSHPKC 482
Qy  816 ----------------SGNSLSH------------------SAVKSLCKTLRRPCLLETL 842
Db  483 KLQKLIFRSVDFLNGCQDFTFLASNKKYTHLDLKETDLGVNGLKTLCEALKCKGCKLRVL 542
Qy  843 RLAGCGLITAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLV 902
Db  543 RLASCDLNVARCQKLSNALQTNRSLVFLNLSLNNLSNDGVKSLCEVLENPNSSLERLAIM 602
Qy  903 SCGLTSDCCQDLASVLSASPSLKEDLLQQNNIDDDVGVRLLCEGLRHPACKLIRIG 957
Db  603 SCVLTSKACGDLASVLVNNSNLWSLDLGHNILDDAGLNILCDALRNPNCHVQRLG 657
```

RESULT 13

```
ID   Q8IWF5;          PRELIMINARY;     PRT;   953 AA.
AC   Q8IWF5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Caspase recruitment domain family, member 4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph;
RA   Strausberg R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC040339; AAH40339.1; -.
SQ   SEQUENCE   953 AA;  107671 MW;  0A9DF167BE87E21A CRC64;
```

Query Match   5.2%;   Score 393.5;   DB 4;   Length 953;
Best Local Similarity   24.0%;   Pred. No. 1.8e-20;
Matches 214;   Conservative 131;   Mismatches 336;   Indels 209;   Gaps 38;

```
Db  306 ETYPFSSCCFTDKEQLKNALLFVQNNAVLYKACQVPGICWVVCSWLKKKMARGQEVSETPSN 365
Qy  547 TTTLCIHYLAQALQAQPLGPQ------LRDLCSLAAEGIWQKKTLFSPDDLRKHGLDG 598
Db  366 STDIFTAYVSTFLPTDGNGDSSELTRHKVLKSLCSLAAEGMRHQRLLFEEEVLRKHGLDG 425
Qy  599 AIISTFLK---------MGILQEHPIPLSYSFIHLCFQEFFAAMSYVLED-----EKGRGKHSN 648
Db  426 PSLTAFLNCIDYRAGLGIKK-----FVSFRHISFQEFFYAMSFLVKEDQSQSGEATHKE 479
Qy  649 C--IIDLEKTLEAYGIHGLFGASTTRFLLGLLSDEGEREMENIFHCRL--SQGRNLMQWV 704
Db  480 VAKLVDPENHEEV--------TLSLQFLPDMLKTEGTLSLGLKFCFRIAPSVRQDLKHFK 531
Qy  705 PSLQLLLQPHSLESLHCLYETRNKTFLTQ 733
Db  532 EQIEAIXYKRSWDLEFSLYDSKIKK-LTQ 559

Qy  237 VVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDPNQKFTQLLLLQRPHPRSQDPLV 296
Db  120 VVNTDPVSRYTQQLRHHLGRDS-KPVLC---------YAQK--EELLLEEIYMDTIMELV 167
Qy  297 KRSWPDYVEENRGHLIEIRDLFG--PGLDTQEBRIVILQGAAGIGKSTLARQVKEAWGRG 354
Db  168 -----GFSNESLGSLNSLACLLDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATG 222
Qy  355 QL-YGDRFQHVFYPSCRELA----QSKVVSLAELIGKD---GTATPAPIRQILSR-PERLL 406
Db  223 RLDAGVKF--FFHFRCRMFSCFKESDRLCLQDLLFKHYCYPBRDPEEVFAFLLRFPHVAL 280
Qy  407 FILDGVDEPGWVLQEPSSELCLH-------WSQPQPADALLGSLLGKTILPEASFLITA 458
Db  281 FTFDGLD---------ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKLITA 332
Qy  459 RTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCL 518
Db  333 RTGI---EVPRQFLRKKVLLRGFSPSHLRAYARRMFPERALQDRLLSQLEANPNLCSLCS 389
Qy  519 VPWVSWLACTCLMQ--------------QMKRKEKLTLTSKTTTL 550
Db  390 VPLFCWIIFRCFQHFRAAFBEGGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNT-- 446
Qy  551 CLHYLAQALQAQPLGPQLRD-LGSL----AAEGIWQKKTLFSPDDLRKHGLDGAIISTFLK 606
Db  447 --HSPVETLHAG------RDTLCGSLGQVAHRGMEKSLFVFTQEEVQASGLQ--------- 494
Qy  607 MGILQEHP--------IPLSYSFIHLCFQEFFAAMSYVLEDEKGR-----------------GK 645
```

RESULT 14

```
Qy   495  LGFLRALPELGLGPGGDQOSYEFFHLTLQAFFTAFFLVLDDRVGTQELLRFFQEWMPPAGAA  554
                                                                      
Db   646  HSNCI------------IDLEKTLEAYGIHGLFGASTTRFLLGLLSDEGE------  683

Qy   555  TTSCYPPPLPFQCLQGSGPAREDLFKNKD----HFQF--TNLFLCGLLSKAKQLLRHL  607
                                                                      
Db   684  ------REMENIFIHCRLSQGRNLMQWVPSL----QLLLQPHSLEHCLYETRNKTF  730

Qy   608  VPAAALRRKKALWAHLFSSLRGYLKSLIPRVQVESFNQVQAMPTFIWWLRCIYETQ----  663
                                                                      
Db   731  LTQVMAHFEEMGMCVETDMELLLVCTFC----IKFSRH-VKKLQL------IEGR  773

Qy   664  -SOKVGQLAARGICANY-LKLTYCNACSADCSALSFVLHHFPKRLALDLDNNNLNDYGVR  721
                                                                      
Db   774  QHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLR  833

Qy   722  ELQPCFSRLTVIRLSVNQITDGGVKVLSEELTKYKILVTYLGLYNNQITDVGARVTKILD  781
                                                                      
Db   834  RPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLITDAGAXHLCQRLR-QP  892

Qy   782  ECKGLTH-LKLGLKNKITSEGGKYLALAVRKNSKSISEVGMWGNQVGDEGAXAFARALRNHP  840
                                                                      
Db   893  SCKLQRLQLVSCGLTSDCCQDIASVLSASPSLKELDLQQNNLDDVGVRLLCEGLR-HPAC  951

Qy   841  S--LTTLSLASNGISTEGGKSLARALQQNTSLEIILWLTQNELNDEVAESLAEMLKVNQTL  898
                                                                      
Db   952  -----LSDEMR---------QELRALEQEK  978

Qy   899  KHLWLIQNQITAKGTAQLADALQSNTGITEICLNGNLIKPEEAKYYEDEK  948
                                                                      
Db   899  (continued)
```

ID   Q8K3Z0   PRELIMINARY;   PRT;   1020 AA.
AC   Q8K3Z0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CARD15.
GN   CARD15.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Iwanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,
RA   Baugh M.M., Suing C., Rosenbaum J.T.;
RT   "Cloning, sequencing and expression analysis of the murine Nod2/Card15
RT   gene.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF520774; AAM76073.1; -.
DR   MGD; MGI:2429397; Card15.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR001130; TatD_DNase.
DR   PROSITE; PS50209; CARD; 2.
DR   PROSITE; PS50053; LRR_RI; 1.
DR   PROSITE; PS50837; NACHT; 1.
DR   PROSITE; PS01137; TATD; 1.
SQ   SEQUENCE 1020 AA; 113561 MW; 255049D5EC?70FBB CRC64;

Query Match   5.2%; Score 392; DB 11; Length 1020;
Best Local Similarity   25.2%;
Matches 186; Conservative 122; Mismatches 316; Indels 114; Gaps 28;

```
Qy   311  LIEIRDLFGP-GLDTQEPRIVILQGEAAGIGKSTLARQVEKEAWGRQOLYGDRFQHVFYFSC  369
                                                                      
Db   255  ILGLEDLFDTHGHLNRDADTILVVGBEAGSGKSTLLQRLHLLWATGRSF-QEFLFIPPPSC  313
```

RESULT 15

```
Qy   370  REL-AQSKVUSLAELIGK-----DGTATPAPIRQILSRPERLLLFILDGVDEPGWVLQEPSS  424
                                                                      
Db   314  RQLQCVAKPLSLRTLLPEHCCWPDVAQDDVFQFLLDHPDRVLLTFDGLDEFKRFTDRER  373

Qy   425  ELCLHWS--QPQPADALLGSLLGKTIIPEASFLITARTLALQNLIPSLEQARWEVLGFS  482
                                                                      
Db   374  ----HCSPIDPTSVQTLLFNLLQGNLLKNACKVLTSRPDAVSALLRKFVRTE-LQLKGFS  428

Qy   483  ESSRKEYFRYFTDERQAIRAFRLVKSNKELWALLCLVPWVSWLACTCLMQQMKREKLTL  542
                                                                      
Db   429  EEGIQLYLRKHREPGVADRLLQLIQATSALHGLCHLPVFSWMVVSRCHRELLLQNRGPFT  488

Qy   543  TSKTTTTLCL-HYLAQALQQP-----QLRDLCSLAAEGIWQKKTLF  586
                                                                      
Db   489  TSTDMYLLILQHFL--LHASPPDSSPLGIGPGLLQSRLSTLLHGLAMSCYVF  545

Qy   587  SPDDLRKHGLDGAIISTFLKMGIL--YSFIHLCFQOEFFAAMSYVLEDEK  641
                                                                      
Db   546  SAQOLQAAQVDSDDIS---LGPLVRAQSSVPGSKAPLEFLHITFQCFFAAFYLAVSADT  601

Qy   642  ----GRGKHSNCI---IDLEKTLEAYGI-----HGLFGASTTRFLL  675
                                                                      
Db   602  SVASLKHLFSCGRLGSSLLGRLLPNLCIQGSRVKKGSEAALLQKAEPHNL--QITRAFLA  659

Qy   676  GLLSDEGEREMENIFIHCRLSQGRNLMQWVPSLQLLLQPHSL------ESLH  720
                                                                      
Db   660  GLLS---QQOHRDLIAAACQVSE-RVLLQRQARARSCL-AHSLREHFHSIPPAVPGETKSMH  714

Qy   721  C---------LYETRNKTFLTQVMAHFE----EMGMCVETDMELLLVCTFCIKFSRHVKK  766
                                                                      
Db   715  AMPGFIWLIRSLYEMQEEQLAQEAVRRLDIGHLKLTFCRVGPAECAALAFVLQHLQRPVA  774

Qy   767  LQL-------IEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNS  819
                                                                      
Db   775  LQLDYNSVGDVGVEQLRPCLGVCTALYLRDNNISDRGARTLVECALRCEQLQKLALFNNK  834

Qy   820  LSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDLARGLRANQTLTELDLSFNVLTD  879
                                                                      
Db   835  LTDACACSMAKLLAHKQNFL-SLRVGNNHITAAGAEVLAQGLKSNTSLKFLGFWGNSVGD  893

Qy   880  AGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDIASVLSASPSLKELDLQQNNLDDVGV  939
                                                                      
Db   894  KGTQALAEVVADHQ-NLKWLSLVGNNIGSMGAEALALMLEKNKSLBELCLBENHICDEGV  952

Qy   940  RLLCEGL-RHPACKLIRL  956
                                                                      
Db   953  YSLAEGLKRNSTLKFLKL  970
```

ID   Q9BY26   PRELIMINARY;   PRT;   287 AA.
AC   Q9BY26;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leucine-rich-repeat protein RNO2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shami P.J., Kanai N., Wang L.Y., Vreeke T.M., Parker C.J.;
RT   "Identification and characterization of a novel gene that is
RT   upregulated in leukemia cells by nitric oxide.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF231021; AAK14942.1; -.
DR   HSSP; P13489; 1A4Y.
DR   InterPro; IPR007091; LRR_RNinh.
DR   PROSITE; PS50503; LRR_RI; 2.
SQ   SEQUENCE 287 AA; 31765 MW; BD3816C325SB2F9E CRC64;

```
Query Match        5.1%;    Score 383;  DB 4;   Length 287;
Best Local Similarity   43.7%;   Pred. No. 1.7e-20;
Matches  93; Conservative  24; Mismatches  94;  Indels   2;  Gaps   1;

Qy    769 LIEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSL 828
          ||  |  |  |     :   :                  : |||:|| :|
Db     73 LCEGLRHPQ--CRLQMIQLRKCQLESGACQGMMASVLGTNPHLVELDLTGNALEBDLGLRLL 130

Qy    829 CKTLRRPRCLLETLRLAGGGLTAEDCKDLAFGLRRNQTLTELDLLSPNVLTDAGAKHLCQR 888
          |:  |||   :||  | ||  | ||||:|||:||||||||  ||:
Db    131 CQGLRHPVCRLRTLWLKICRLTAAACDELASTLSVNQSLRELDLSINELGDLGVLLLCEG 190

Qy    889 LRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRH 948
          ||  |:||| |||||:  |:|  |||:|:   |:|:|||||  :|||  :||||||: :
Db    191 LRHPTCKLQTLRLDSCGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLCKRLSH 250

Qy    949 PACKLIRLGLDQTTLSDBMRQELRALBQEKPQL 981
          | ||| |  | :       |   |
Db    251 PGCKLRVLWLFGMDLNKMTHSRLAALRVTKPYL 283
```

Search completed: January 29, 2004, 13:48:11
Job time : 84.1934 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 29, 2004, 13:11:31 ; Search time 28.1576 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title:
Perfect score: 7534
Sequence: 1 MAGGAWGRLACYLEFLKKEE.........HLIMELWEKGSKKGLLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	388.5	5.2	953	4	US-09-099-041A-8	Sequence 8, Appli
2	388.5	5.2	953	4	US-09-245-281-8	Sequence 8, Appli
3	388.5	5.2	953	4	US-09-207-359B-8	Sequence 8, Appli
4	388.5	5.2	953	4	US-09-340-620A-8	Sequence 8, Appli
5	363.5	4.8	456	2	US-08-910-731-8	Sequence 8, Appli
6	361.5	4.8	456	2	US-08-910-731-4	Sequence 4, Appli
7	361.5	4.8	456	2	US-08-795-395-4	Sequence 4, Appli
8	357	4.7	461	2	US-08-910-731-6	Sequence 6, Appli
9	354.5	4.7	456	2	US-08-910-731-2	Sequence 2, Appli
10	354.5	4.7	456	2	US-08-795-395-2	Sequence 2, Appli
11	348	4.6	456	2	US-09-245-281-43	Sequence 43, Appli
12	348	4.6	953	4	US-09-207-359B-43	Sequence 43, Appli
13	348	4.6	953	2	US-09-340-620A-43	Sequence 43, Appli
14	345	4.6	1130	2	US-08-519-547A-6	Sequence 6, Appli
15	326.5	4.3	966	4	US-09-207-359B-47	Sequence 47, Appli
16	253.5	3.4	193	4	US-09-340-620A-61	Sequence 61, Appli
17	250.5	3.3	195	4	US-09-340-620A-49	Sequence 49, Appli
18	218.5	2.9	490	4	US-09-099-041A-26	Sequence 26, Appli
19	218.5	2.9	490	4	US-09-245-281-26	Sequence 26, Appli
20	218.5	2.9	490	4	US-09-207-359B-26	Sequence 26, Appli
21	218.5	2.9	490	4	US-09-340-620A-26	Sequence 26, Appli
22	216	2.9	71	4	US-09-340-620A-58	Sequence 58, Appli
23	209	2.8	70	4	US-09-340-620A-66	Sequence 66, Appli
24	209	2.8	70	4	US-09-340-620A-57	Sequence 57, Appli
25	198	2.6	1151	4	US-08-836-134-23	Sequence 23, Appl
26	198	2.6	1151	4	US-09-493-784-23	Sequence 23, Appl
27	198	2.6	1232	3	US-08-836-134-2	Sequence 2, Appli
28	198	2.6	1232	3	US-09-493-784-2	Sequence 2, Appli
29	179.5	2.4	483	4	US-09-904-615-154	Sequence 154, App
30	175.5	2.3	200	4	US-09-099-041A-11	Sequence 11, Appl
31	175.5	2.3	200	4	US-09-245-281-11	Sequence 11, Appl
32	175.5	2.3	200	4	US-09-207-359B-11	Sequence 11, Appl
33	175.5	2.3	200	4	US-09-340-620A-11	Sequence 11, Appl
34	165	2.2	1466	4	US-09-252-991A-30085	Sequence 30085, A
35	147.5	2.0	793	3	US-09-012-710-10	Sequence 10, Appl
36	147.5	2.0	793	3	US-09-556-273-10	Sequence 10, Appl
37	147	2.0	2482	1	US-08-328-254-6	Sequence 6, Appli
38	143.5	1.9	4302	4	US-09-052-469-8	Sequence 8, Appli
39	143.5	1.9	4302	4	US-09-422-582-8	Sequence 8, Appli
40	143.5	1.9	4339	4	US-09-052-469-6	Sequence 6, Appli
41	143.5	1.9	4339	4	US-09-422-582-6	Sequence 6, Appli
42	142.5	1.9	4302	2	US-08-422-582-6	Sequence 6, Appli
43	142.5	1.9	4303	2	US-08-658-136-5	Sequence 5, Appli
44	138.5	1.8	794	1	US-08-460-751-2	Sequence 2, Appli
44	138.5	1.8	794	1	US-08-393-333-2	Sequence 2, Appl
45	138.5	1.8	794	3	US-09-087-465-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-099-041A-8
; Sequence 8, Application US/09099041A-8
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
Matches 217; Pred. No. 3.7e-28;

```
Qy  237 VVGTPPQAHTSLQPHHPWPEPSVRESLCSTWPWKDNEDFNQKFTQLLLLQRPHPRSQDPLV 296
            |:|                                        :||||
Db  120 VVNTDPVSRYTQQLRHHLGRDS-KPVLC----------YAQX--EELLLEEIYMDTIMELV 167

Qy  297 KRSWPDYVEENRGHLIBIRDLFG--PGLDTQBPRIVILQGAAGIGKSTLARQVKEAWGRG 354
            :  |             ||  |:|  |        |||||:||||||
Db  168 -----GFSNESLGSLNSLACLLDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATG 222

Qy  355 QL-YGDRFQHVFFSCRELA----QSKVVSLAELIGKD---GTATPAPIRQILSR-PERLL 406
            || :  ||:|                :       :|    |  :|  |:|:|  |||||
Db  223 RLDAGVKP--PFHFRCRWFSCFKESDRLCLQDLLFKHYCYPERDPEEVFAFLLRPPHVAL 280

Qy  407 FILDGVDEPGWVLQEPSSELCLH--------WSQPQPADALLGSLLGKTILPEASFLITA 458
            | : | :    :| :||| |||        |  | |||:||||:|||||:| |:|||||
Db  281 FTFDPGLD-------ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKLLTA 332

Qy  459 RTTALQNLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCL 518
            ||               |:|||:|:|||:|||||||            ::|:|:| |||
Db  333 RTGI----EVPRQFLRKKVLLRGPSPSHLRAYARMFPERALQDRLLSQLEANPNICSLCS 389

Qy  519 VPWVSWLACTCLMQQMKRKE---------KLTLTSKTTTTLCLHY------LAQALQAQP 563
            || : |:: :|:||:  :|         |||||||||||||:|||      ||| || |
Db  390 VPLFCWIIFRCFQHFRAAFEGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNTRSP 449

Qy  564 L-----GPQLRD-LCSL---AAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGTLQEHP 614
```

RESULT 2

```
US-09-245-281-8
Sequence 8, Application US/0924528I
Patent No. 6369196

GENERAL INFORMATION:
  APPLICANT: Bertin, John
  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
  TITLE OF INVENTION: AND USES THEREOF
  FILE REFERENCE: 07334/118001
  CURRENT APPLICATION NUMBER: US/09/245,281
  CURRENT FILING DATE: 1999-02-05
  EARLIER APPLICATION NUMBER: US 09/207,359
  EARLIER FILING DATE: 1998-12-08
  EARLIER APPLICATION NUMBER: US 09/099,041
  EARLIER FILING DATE: 1998-06-17
  EARLIER APPLICATION NUMBER: US 09/019,942
  EARLIER FILING DATE: 1998-02-06
  NUMBER OF SEQ ID NOS: 44
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 8
  LENGTH: 953
  TYPE: PRT
  ORGANISM: Homo sapiens

Query Match   24.6%;   Score 388.5;   DB 4;   Length 953;
Best Local Similarity  24.6%;
Pred. No. 3.7e-28;
Matches 217;  Conservative 129;  Mismatches 343;  Indels 193;  Gaps 39;

Db  450  VETLHAG------RDTLCSLGQVAHRGMEKSLFVFTQBEEVQASGLQ-----ERDMQLGFLRALP  502
Qy  615  -IPLSYSPIHLCPQEFFRAAMSYVLBDEKGR------GKHSNCI---  650
Db  503  ELGPGGDQOSYEFFHLTLQAFFTAFFLVLDDRVGTQELLRFFQEWMPPAGAATTSCYPPF  562
Qy  651  ------IDLEKTLEAYGIHGLFGASTTRFLLGLLSDEGE------R  684
Db  563  LPFQCLQGSGPAREDLFQNKD----HFQF---TNLFLCGLLSKAKQKLLRHLVPAAALRR  615
Qy  685  EMENIFHCRLSQGRNLMQWVPSL------QLLLQPHSLESLHCLYETRNKTFLTQVMAHF  738
Db  616  KRKALWAHLFSSLRGYLIKSLPRVQVESFNQVQAMPTFIWMLRCIYETQ----SQKVGQL  670
Qy  739  ------IKFSRH--VKKLQL------IEGRQHRSTWSP  781
Db  671  AARGICANY-LKLTYCNACSADCSALSFVLIHFPKRLALDLDNNNLINDYGVRELQPCFSR  729
Qy  782  TMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLET  841
Db  730  LTVLRLSVNQITDGGVKVLSEELTKYKIVTYLGLYNNQITDVGARYVTKILDECKGLTH-  788
Qy  842  LRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLR-QPSCKLQRLQ  900
Db  789  LKLGKNKITSEGGKYLALAVRSNKSISEVGMWGNQVGDEGAKAFAEALRNHPS----LTTLS  846
Qy  901  LVSCGLITSDCCQDLASVLSASPSIKELDLQQNNLDDVGVRLLCEGLR-HPACKLIRLGLD  959
Db  847  LASNGISTEGGKSLARALQONTSLEILMLTQNELNDEVAESLAEMLKVNQTLKHLWLIQN  906
Qy  960  -----LSDEMR------QELRALEQEK  978
Db  907  QITAKGTAQLADALQSNTGITEICLNGNLIKPEBAKYEDEK  948
```

RESULT 3

```
US-09-207-359B-8
Sequence 8, Application US/09207359B
Patent No. 6469140

GENERAL INFORMATION:
  APPLICANT: Bertin, John
  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
  TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
  FILE REFERENCE: 07334-112001
  CURRENT APPLICATION NUMBER: US/09/207,359B
  CURRENT FILING DATE: 1998-12-08
  PRIOR APPLICATION NUMBER: US 09/099,041
  PRIOR FILING DATE: 1998-06-17
  PRIOR APPLICATION NUMBER: US 09/019,942
  PRIOR FILING DATE: 1998-02-06
  NUMBER OF SEQ ID NOS: 47
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 8
  LENGTH: 953
  TYPE: PRT

Qy  355  QL-YGDRFQHVFYFSCRELA---QSKVVSLAELIGKD---GTATPAPIRQILSR-PERLL  406
Db  223  RLDAGVKF--FFHFRCRMFSCFKFKBSDRLCLQDDLLFKHYCYPERDPBEVFAPLLRFPHVAL  280
Qy  407  FILDGVDEPGWVLQEPSSELCLH------WSQPQPADALLGSLLGKITLIPEASFLITA  458
Db  281  FTPDGLD-----ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKLITA  332
Qy  459  RTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCL  518
Db  333  RTGI---EVPRQFLRKKVLLRGFSPSHLRAYAARRMFPPERALQDRLLSQEANPNLCSLCS  389
Qy  519  VPWVSWLACTCLMQQMKRKE------KLTLISKTTTTLCLHY-----LAQALQAQP  563
Db  390  VPLFCWIIFRCFQHFRAAFEGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNTRSP  449
Qy  564  L------GPQLRD-LCSL-----AAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHP  614
Db  450  VETLHAG------RDTLCSLGQVAHRGMEKSLFVFTQBEEVQASGLQ---ERDMQLGFLRALP  502
Qy  615  ------IPLSYSPIHLCPQEFFRAAMSYVLBDEKGR------GKHSNCI--  650
Db  503  ELGPGGDQOSYEFFHLTLQAFFTAFFLVLDDRVGTQELLRFFQEWMPPAGAATTSCYPPF  562
Qy  651  ------IDLEKTLEAYGIHGLFGASTTRFLLGLLSDEGE------R  684
Db  563  LPFQCLQGSGPAREDLFQNKD---HFQF---TNLFLCGLLSKAKQKLLRHLVPAAALRR  615
Qy  685  EMENIFHCRLSQGRNLMQWVPSL------QLLLQPHSLESLHCLYETRNKTFLTQVMAHF  738
Db  616  KRKALWAHLFSSLRGYLIKSLPRVQVESFNQVQAMPTFIWMLRCIYETQ----SQKVGQL  670
Qy  739  ------IKFSRH--VKKLQL------IEGRQHRSTWSP  781
Db  671  AARGICANY-LKLTYCNACSADCSALSFVLIHFPKRLALDLDNNNLINDYGVRELQPCFSR  729
Qy  782  TMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLET  841
Db  730  LTVLRLSVNQITDGGVKVLSEELTKYKIVTYLGLYNNQITDVGARYVTKILDECKGLTH-  788
Qy  842  LRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLR-QPSCKLQRLQ  900
Db  789  LKLGKNKITSEGGKYLALAVRSNKSISEVGMWGNQVGDEGAKAFAEALRNHPS--LTTLS  846
Qy  901  LVSCGLITSDCCQDLASVLSASPSIKELDLQQNNLDDVGVRLLCEGLR-HPACKLIRLGLD  959
Db  847  LASNGISTEGGKSLARALQONTSLEILMLTQNELNDEVAESLAEMLKVNQTLKHLWLION  906
Qy  960  -----LSDEMR------QELRALEQEK  978
Db  907  QITAKGTAQLADALQSNTGITEICLNGNLIKPEBAKYEDEK  948
```

Query Match   Best Local Similarity 24.6%;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
Score 388.5; DB 4; Length 953; Pred. No. 3.7e-28;

```
Qy  237 VVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFPNQKFTQLLLLQRPHPRSQDPLV 296
        ::  | |   |  |:|| |                   :    ::
Db  120 VVNTDPVSRYTQQLRHHLGRDS-KPVLC------------------YAQK--EELLLEEIYMDTIMELV 167

Qy  297 KRSWPDYVEENRGHLIEIRDLFG--PGLDTQEPRIVILQGAAAGIGKSTLARQVKEAWGRG 354
        :    :: |:|  :: | | :     |||:|:    |:||:|:|||||::| ||:|: |:
Db  168 -----GFSNESLGSLNSLACLLDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATG 222

Qy  355 QL-YGDRFQHVFYFSCRELA---QSKVVSLAELIGKD---GTATPAPIRQILSR-PERLL 406
        |: | ||||||||||:  |     :||: | | ||     ||||| ||||||:| |:|||
Db  223 RLDAGVKF--FFHFRCRMFSCFKESDRLCLQDLLFKHYCYPERDPEEVFAFLLRPFHVAL 280

Qy  407 PILDGVDEPGWVLQEPSSELCIH-------WSQPQPADALLGSLLGKTILPEASFLITA 458
        |  ||:|   |||||||||||        ||||||||||||||||||||||||||||:
Db  281 FTFPDGLD------ELHSDLDLSRVPDSSCPWEPAHPL-VLIANLLSGKLLKGASKLLTA 332

Qy  459 RTTALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCL 518
        || |: ||:|||||||:|||||||| :|||||:|||||||:|||||||||:||:||:|||
Db  333 RTGI--EVPRQFLRKKVLLRGFSPSHLRAYARRMFPERALQDRLLSQLEANPNLCSLCS 389

Qy  519 VPWVSWLACTCLMQQMKRKE----------KLTLTSKTTTTLCLHY------LAQALQAQP 563
        || |  |   |: |: |             ||||||||||||||||      |||||||||
Db  390 VPLFCWIIFRCFQHFRAAFEGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNTRSP 449

Qy  564 L-----GPQLRD-LCSL---AAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHP 614
        |     |||||| |  |   |||||:|||||||||:|||:|||||||||||||||||||||
Db  450 VETLHAG---RDTLCSLGQVAHRGMEKSLFVFTQEEVQASGLQ---ERDMQLGFLRALP 502

Qy  615 ------IPLSYSFIHLCFQEFFAAMSYVLEDEKGR------------GKHSNCI-- 650
              |||||||||:||||||||||||||||||||             ||||||:
Db  503 ELGPGGDQQSYEFFHLTLQAPFTAFFLVLDDRVGTQELLRFFQEWMPPAGAATTSCYPPF 562

Qy  651 ------IDLEKTLEAYGIHGLFGASTTRFLLGLLSDEGB------R 684
              |||||||||:||||||||||||||||||||||||:      |
Db  563 LPPQCLQGSGPAREDLFKND----HFQF--TNLFLCGLLSKAKQKLLRHLVPAAALRR 615

Qy  685 EMENIFHCRLSQGRNLMQWVPSL-----QLLLQPHSLESLHCLYETRNKTFLTQVMAHF 738
        |||||||||||||||||||||||     |||||||||||||||||||||||||||||||
Db  616 KRKALWAHLFSSLRGYLKSLPRVQVESFNQVQAMPTFIWMLRCIYETQ----SQKVGQL 670

Qy  739 EEMGMCVETDMELLVCTFC----------IKFSRH--VKKLQL---IEGRQHRSTWSP 781
        |||||||||||||||||||          ||||||  ||||||   |||||||||||||
Db  671 AARGICANY-LKLTYCNACSADCSALSFVLHHFPKRLALDLDNNNLNDYGVRELQPCFSR 729

Qy  782 TMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDISGNSLSHSAVKSLCKTLRRPCLLET 841
        :|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  730 LTVLRLSVNQITDGGVKVLSEELTKYKIVTYLGLYNNQITDVGARYVTKILDECKGLTH- 788

Qy  842 LRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLR-QPSCKLQRLQ 900
        |:|  | |||  ||:||| |||||:||||||||||||||||||||||||  | |||||:|
Db  789 LKLGKNKITSEGGKYLALAVRNSKSISEVGMWGNQVGDEGAKAFAEALRNHPS--LTTLS 846

Qy  901 LVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLR-HPACKIRLGLD 959
        |:| |:|::   :|||:::|:|||||||| ||| |||| : | |  |   ||| | | |
Db  847 LASNGISTEGGKSLARALQQNTSLEILWLTQNELNDVABSLAEMLKVNQTLKHLWLIQN 906

Qy  960 QTT-----------LSDEMR----------QELRALEQEK 978
        |          | :|:|          |||||:||||
Db  907 QITARGTAQLADALQSNTGITEICLNGNLIKPEEAKVEDEK 948
```

RESULT 5
US-08-910-731-8
; Sequence 8, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-910-731-8

Query Match   4.8%; Score 363.5; DB 2; Length 456;
Best Local Similarity   25.5%; Pred. No. 2.7e-26;
Matches 144; Conservative 54; Mismatches 201; Indels 165; Gaps 15;

Qy   470 LEQARWEVL------------------GFSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWA 515
                |||||||| |||||||||| |||| ||||
Db   10 LSDARWTELLPLLQQYEVVRLDDCGLTEEHCK------DIGSALRA-----NPSLTE 55

Qy   516 LCLVPWVSWLACTCLMQQMKRKEKLTLTSKTTTLCLHYIAQALQAQPLGPQLRDL--CS 573
                ||| ||| ||| ||
US-08-910-731-8

Qy   842 LRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLR-QPSCKLQRLQ 900
           |||| ||| || ||| || |||| ||||| ||||||||||| | |||||||
Db   789 LKLGRNKITSEGGKYLALAVKQNSKSISEVGWGMNQVGDEGAKAFAEALRNHPS--LTTLS 846

Qy   901 LVSCGLITSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLR-HPACKLRLGLD 959
           || ||| | | || | ||| |||| ||| |||||| ||||| || |||||||| ||
Db   847 LASNGISTEGGKSLARALQONTSLEILMLTQNELNDEVAESLAEMLKVNQTLKHLWLIQN 906

Qy   960 QTT-------------LSDEMR------------------QELRALEQEK 978
                         | || ||||||||
Db   907 QITAKGTAQLADALQSNTGITEICLNGNLIKPEEAKVIEDEK 948

RESULT 6
US-08-910-731-4
; Sequence 4, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003

Db   56 LCL-------------------RTNELGDAG--------VHLVLQGLLQSPTCKILSLQNCS 91
                                                      |
Qy   574 LAAEGIWQKKTLFSPDDLRKHGLDGAISTFLKQMGILQEHPIPLSYSFHLCFQEFFAAM 633
                                                                      WI 258
Db   92 LTEAGC--------------GVLPSTIRSLPTLRE-------------LHLS---- 116

Qy   634 SYVLEDEKGRGKHSNCIIDLEKTLEAYGIHGLPGASTRFLL-GLLSDEGEREMENIFHC 692
                                                                      |
Db   117 --------------------DNPLGDAGLRLLCEGLLDPQCHLEKLQLEYC 147

Qy   693 RLSQGRNLMQWVPSLQLLLQPHSLESL--------------------HCLYETRNKTF 730

Qy   148 RLTAA-----SCEPLASVLRATRALKELTVSNNDIGEAGARVLGQGLADSACQLETLRLEN 203

Qy   731 LTQVWAHFEEMGMCVETDMELLVCTFCIKFSRHYKKLQLIEGRQHRSTWSPTMVVLFRWV 790

Qy   204 CGLTPANCKDLCGIVASQASLRELALGSNKLGPDVGWMAELCPGLLHPSSRLRTL----WI 258

Qy   791 PVTDAYWQI--------LFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRRCLLET 841

Qy   842 LRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQL 901

Db   259 ----WECGITAKGCGDLCRVLRAKESLKELSLAGNELGDEARLLCETLLEPGCQOLES 312

Db   313 LWYKSCSFTAACCPHFSSVLAQNRFLLELQISNNRLEDAGVRELCQGLGQPGSVLRVLML 372

Db   373 ADCDVSDSCSCSSLAATLLANHSLRELDLSNNCLGDAGILQLVESVRQPGCLLEQLVLYDI 432

Db   902 VSCGLITSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLRLGLDQT 961

Db   962 TLSDEMRQELRALEQEKPQLLIFS 985

Db   433 YWSEEMEDRLQALEKDKPSLRVIS 456

Left column:

```
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 202-371-2600
;    TELEFAX: 202-371-2540
;  INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 456 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein

US-08-910-731-4

Query Match      4.8%;  Score 361.5;  DB 2;  Length 456;
Best Local Similarity  35.5%;  Pred. No. 4.2e-26;
Matches  89;  Conservative  32;  Mismatches  73;  Indels  57;  Gaps  2;

Qy  792 VTDAYWQILFSVLKVTRNIKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTA 851
         :|:|||||:::|::|||||:|||:::|||||::::::|::|||::||:.|:||:||||
Db  206 ITSANCKDLCDVVASKASLQELDLGSNKLGNTGIAALCSGLLLPSCRLRTLWLMDCDVTA 265

Qy  852 EDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTS--- 908
         |.||||..:|.|:.:|.||.|:.:|.:..|:||.||:..||.|:|.||.::.|.||
Db  266 EGCKDLCRVLRAKQSLKELSLAGNELKDEGAQLLCESLLEPGCQLESLWVKTCSLTAASC 325

Qy  909 --------------------------------------------DC------CQDL 914
                                                        .|      .|||
Db  326 PHFCSVLTKNSSLFELQMSSNPLGDSGVVELCKALGYPDTVLRVLWLGDCDVTDSGCSSL 385

Qy  915 ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHIPACKLIRLGLDQTTLSDEMRQELRAL 974
         |:||.|.|||:||||..||..:..|:||.||||:|.|:||:||||.||:||:|:::||||
Db  386 ATVLLANRSLRELDLSNNCMGDNGVLQLLESLKQPSCILQQLVLYDIYWTDEVEDQLRAL 445

Qy  975 EQEKPQLLIFS 985
         |:|:||:|:|S
Db  446 EEERPSLRIIS 456


RESULT 7
US-08-795-395-4
Sequence 4, Application US/0875395
Patent No. 5965399
GENERAL INFORMATION:
  APPLICANT: CHATTERJEE, DEB K.
  APPLICANT: SHANDILYA, HARINI
  TITLE OF INVENTION: Cloning and Expression of Rat Liver and
  TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
  NUMBER OF SEQUENCES: 8
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
    STREET: 1100 NEW YORK AVE., N.W., SUITE 600
    CITY: WASHINGTON
    STATE: D.C.
    COUNTRY: USA
    ZIP: 20005-3934
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/795,395
    FILING DATE: 04-FEB-1997
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/024,057
    FILING DATE: 16-AUG-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: ESMOND, ROBERT W.
    REGISTRATION NUMBER: 32,893
    REFERENCE/DOCKET NUMBER: 0942.3440002
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-371-2600
    TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
```

Right column:

```
;  SEQUENCE CHARACTERISTICS:
;    LENGTH: 456 amino acids
;    TYPE: amino acid
;    TOPOLOGY: linear
;  MOLECULE TYPE: protein

US-08-795-395-4

Query Match      4.8%;  Score 361.5;  DB 2;  Length 456;
Best Local Similarity  35.5%;  Pred. No. 4.2e-26;
Matches  89;  Conservative  32;  Mismatches  73;  Indels  57;  Gaps  2;

Qy  792 VTDAYWQILFSVLKVTRNIKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTA 851
         :|:|||||:::|::|||||:|||:::|||||::::::|::|||::||:.|:||:||||
Db  206 ITSANCKDLCDVVASKASLQELDLGSNKLGNTGIAALCSGLLLPSCRLRTLWLMDCDVTA 265

Qy  852 EDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTS--- 908
         |.||||..:|.|:.:|.||.|:.:|.:..|:||.||:..||.|:|.||.::.|.||
Db  266 EGCKDLCRVLRAKQSLKELSLAGNELKDEGAQLLCESLLEPGCQLESLWVKTCSLTAASC 325

Qy  909 --------------------------------------------DC------CQDL 914
                                                        .|      .|||
Db  326 PHFCSVLTKNSSLFELQMSSNPLGDSGVVELCKALGYPDTVLRVLWLGDCDVTDSGCSSL 385

Qy  915 ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHIPACKLIRLGLDQTTLSDEMRQELRAL 974
         |:||.|.|||:||||..||..:..|:||.||||:|.|:||:||||.||:||:|:::||||
Db  386 ATVLLANRSLRELDLSNNCMGDNGVLQLLESLKQPSCILQQLVLYDIYWTDEVEDQLRAL 445

Qy  975 EQEKPQLLIFS 985
         |:|:||:|:|S
Db  446 EEERPSLRIIS 456


RESULT 8
US-08-910-731-6
Sequence 6, Application US/08910731
Patent No. 5932440
GENERAL INFORMATION:
  APPLICANT: CHATTERJEE, DEB K.
  APPLICANT: SHANDILYA, HARINI
  TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
  NUMBER OF SEQUENCES: 16
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
    STREET: 1100 NEW YORK AVE., N.W., SUITE 600
    CITY: WASHINGTON
    STATE: D.C.
    COUNTRY: USA
    ZIP: 20005-3934
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/910,731
    FILING DATE: (Herewith)
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/794,546
    FILING DATE: 03-FEB-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/795,395
    FILING DATE: 04-FEB-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/024,057
    FILING DATE: 16-AUG-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: ESMOND, ROBERT W.
    REGISTRATION NUMBER: 32,893
    REFERENCE/DOCKET NUMBER: 0942.3440003
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-371-2600
    TELEFAX: 202-371-2540
```

```
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 461 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: not relevant
;     MOLECULE TYPE: peptide

US-08-910-731-6

Query Match             4.7%;  Score 357;  DB 2;  Length 461;
Best Local Similarity  37.9%;  Pred. No. 1.2e-25;
Matches   86;  Conservative   30;  Mismatches   91;  Indels   20;  Gaps    3;

Qy    768 QLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQI-----------LFSVLKVTRNLKELDLSGN 818
              :|:|   :      |  :           |              |  | || ::|||:|
Db    246 ELCPGLLIPSSRLRTL-----WI------WECGITAKGCGDLCRVLRAKEBLKELSLAGN 294

Qy    819 SLSHSAVKSLCKTLRRPRCLLETLRLAGGGLTAEDCKDLAFGLRANQTLTELDLSFNVLT 878
              :|:|| :||   |||  |: | |:| ||| ||||||| ||:|: |    |:|:|: ||
Db    295 ELGDEGARLLCETLLEPGCQOLESLWVKSCSFTAACCPHFSSVLAQNRFLLELQISNNRLE 354

Qy    879 DAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKELDLQQNNLDDVG 938
              |||  |||||||||| ||||||||||||  ::| |: :| ||||:||::|||::||:|
Db    355 DAGVRELCQGLGQPGSVLRVLWLLADCDVSDSSCSSLAATLLANHSLRELDLSNNCIGDAG 414

Qy    939 VRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELRALEQEKPQLLIFS 985
              |||||||| :|:|:  :||:|||||:|||:|:|:|:|:|  :|:::
Db    415 ILCLVESVRQPGCLLEQLVLYDIYWSEEMEDRLQALEKDKPSLRVIS 461

RESULT 9
US-08-910-731-2
; Sequence 2, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
;   APPLICANT: CHATTERJEE, DEB K.
;   APPLICANT: SHANDILYA, HARINI
;   TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
;   NUMBER OF SEQUENCES: 16
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;     STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;     CITY: WASHINGTON
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3934
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/910,731
;     FILING DATE: (herewith)
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/795,395
;     FILING DATE: 04-FEB-1997
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/794,546
;     FILING DATE: 03-FEB-1997
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/024,057
;     FILING DATE: 16-AUG-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: ESMOND, ROBERT W.
;     REGISTRATION NUMBER: 32,893
;     REFERENCE/DOCKET NUMBER: 0942.3440003
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-371-2600
;     TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
```

```
;     LENGTH: 456 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein

US-08-910-731-2

Query Match             4.7%;  Score 354.5;  DB 2;  Length 456;
Best Local Similarity  34.3%;  Pred. No. 2e-25;
Matches  108;  Conservative   35;  Mismatches  101;  Indels   71;  Gaps    8;

Qy    752 LVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSVLKVTRNLK 811
              |:|   |:|    :|          |               |          :::|
Db    127 LLCEGLLDPQCHLEKLQL-----------------------EYCRLTAASCEPLASVLRATRALK 168

Qy    812 ELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGGGLTAEDCKDLAFGLRANQ-TLTEL 870
              |||||| : |  |: :| |||| ||:| | |:|:|| ||||||| |:|:  :  |||:|
Db    169 ELTVSNNDIGEAGARVLGQGLADSACQLETLRLENCGLTPANCKDLC-GIVASQASLREL 227

Qy    871 DLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLLASVLSASPSLKELDLQ 930
              |||| :|| | |  |||||||||| |||||||||||||  | ||||: :| |||:||:||
Db    228 DLGSNGLGDAGIAELCPGLLSPASRLKTLWLWEDCITASGCRDLCRVLQAKETLKELSLA 287

Qy    931 QNNLDDVGVRLLCEGLRHPACK----------------------LIRLGLDQTT 962
              |||:||:| ||||||||:||||                      |:|:||:||
Db    288 GNKLGDEGARLLCESLLQPGCQOLESLWVKSCSLTAACCQHVSLMLITQNKHLLELQLSSNK 347

Qy    963 LSDEMRQEL-RALEQEKPQLLIFSRRKPSVMTFTEGLIDTGEMSNS-TSSLKRQRLGSERA 1020
              || |:||:  :|:|:|:| ::| :: | | :|:|::|:   |  :     :|:|::| :
Db    348 LGDSGIQELCQALSQPGTTLRVLC-------------------------LGDCEVTNSGCSSL 401

Qy   1021 ASHVAQANLKLLDVS 1035
              | : ::| :|||:|
Db    387 SLLLANRSLRELDLS 401
```

```
RESULT 10
US-08-795-395-2
; Sequence 2, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
;   APPLICANT: CHATTERJEE, DEB K.
;   APPLICANT: SHANDILYA, HARINI
;   TITLE OF INVENTION: Cloning and Expression of Rat Liver and
;   TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;     STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;     CITY: WASHINGTON
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3934
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/795,395
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/024,057
;     FILING DATE: 16-AUG-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: ESMOND, ROBERT W.
;     REGISTRATION NUMBER: 32,893
;     REFERENCE/DOCKET NUMBER: 0942.3440002
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-371-2600
;     TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
```

LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-395-2

Query Match   4.7%;   Score 354.5; DB 2; Length 456;
Best Local Similarity   34.3%;   Pred. No. 2e-25;
Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

```
Qy   752 LVCTFCIKFSRHVKKLQLIEBGRQHRSTWSPTMVULFRWVPVTDAYWQILFSVLKVTRNLK 811
Db   108 
Qy       EYCRLTAASCEPLASVLRATRALK 168
Db   127 LICEGLLDPQCHLEKLQL
Qy   812 ELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCDLAFGLRANQ-TLTEL 870
Db   169 ELTVSNNDIGEAGARVLGQGLADSACQLETLRLENCGLTPANCKDLC-GIVASQASLREL 227
Qy   871 DLSFNVLTDAGAXHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQ 930
Db   228 DLGSNGLGDAGIAELCPGLLSPASRLKTLWLWECDITASGCRDLCRVLQAKETLKELSLA 287
Qy   931 QNNLDDVGVRLLCEGLRHPACK 962
Db   288 GNRLGDEGARLLCESLLQPGCQLESLVWKSCSLTAACCQHVSLMLTQNKHLLELQLSSNK 347
Qy   963 LSDEMRQEL-RALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNS-TSSLKRORLGSERA 1020
Db   348 LGDSGIQELCQALSQPGTTLRVLC
Qy  1021 ASHVAQANLKLLDVS 1035
Db       LGDCEVTNSGCSSL----------A 386
Db   387 SLILANRSLRELDLS 401
```

RESULT 11
US-09-245-281-43
Sequence 43, Application US/09245281
Patent No. 6369196
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 07334/118001
CURRENT APPLICATION NUMBER: US/09/245,281
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: US 09/207,359
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: US 09/099,041
EARLIER FILING DATE: 1998-06-17
EARLIER APPLICATION NUMBER: US 09/019,942
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 44
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 43
LENGTH: 953
TYPE: PRT
ORGANISM: Homo sapiens

Query Match   4.6%;   Score 348; DB 4; Length 953;
Best Local Similarity   21.6%;   Pred. No. 3.4e-24;
Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

```
Qy   234 WAAVVGTFPQAHTSLQPHHPWPWBPSVRESLCSTWPWKNEDFPNQKFTQLLLLQRP----- 287
Db   103 WLSEIGFSP------------------SQLIRTKTIVANTDPVSRYTQOLRHQLGRDSKFM 144
Qy   288 --HPRSQDPLVKRSWPD----------YVEENRGHLIEIRDLF--GPGLDTQEPRIVILQGAA 336
Db   145 LCYAQKEDLLLEETYMDTLMGLVGFNNENLGSLGGLDCLLLDHSTGVLNEHGETVFVFGDA 204
Qy   337 GIGKSTLARQVKEAWGRGQLYGDRFQHVYFYFSCRELA---QSKVVSLAELIGKD---GTA 390
```

RESULT 12
US-09-207-359B-43
Sequence 43, Application US/09207359B
Patent No. 6469140
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-112001
CURRENT APPLICATION NUMBER: US/09/207,359B
CURRENT FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSEQ for Windows Version 4.0

```
Db   205 GVGRKSMLLQRLQSLWASGRLTSTA-KFFFHFRCRMFSCFKESDMLSLQDLLFKHFCYPEQ 263
Db   264 DPEEVFSFLLRFPHTALFTFDGLDELHSPDPLSRVPDS--CCPWBPAHPL-VLLANLLSG 320
Db   321 RLLKGAGKLLTARTGV---EVPRQLLRKKVLLRGPSPSHLRAYARRMFPPRTAQEHLLQQ 377
Db   378 LDANPNLCSLCGVPLFCWIIFRCFQHFQTVFEGSSSSQLPDCAVTLTDVFLLVTEVHLNRP 437
Qy   447 TILPEASFLITARTTALQNLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRL 506
Db   438 QPSSLVQRNTRS------PAETLRAG--WRTLHALGEVAHRGTDKSLLFVFGQBEVQASKLQ 490
Qy   507 VKSNKELWALCLVPMVWSWLACTCLMQ--------------------------QMKRK 537
Db   491 ---EGDLQLGFLRALPDVGPEQCGQSYEFFHLTLQAFFTAFFILVADDKVSTRELLRFFRE 546
Qy   598 GAIISTFLKMGILQEHP-----IPLSYSFIHLCFQBFFAAMSYVLEDE------------ 640
Db   547 WTSPGEATSSSCHSSF--------FSFQCLGGRSRLGPDPFRNKDHHQFTNLFVCGLLAKAR 600
Qy   641 ---KGRGKHSNCIIDLEKTLEAYGIHGLFGAS------------------TRFLLGILSDEG 682
Db   601 QKLLRQLVPKAILRRKRKALWAHLFPASLRSYLKSLPRVQSGGFNQVHAMPTFLWMLRCIY 660
Qy   683 EREMENIFHCRLSQCRNLMQW------------VPSL-----QLLLQHSLESLHCLY 723
Db   661 ETQ------SQKVGRLAAARGI----SADYLKLAFCNACSADCSALSFVLHHFHRQL--- 706
Qy   724 ETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEBGRQHRSTWSPTM 783
Db   707 ----------------ALDLDNNLNDYGVQEILQOPCFSR---LTVIR 734
Qy   784 VVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLR 843
Db   735 LSVNQITDTGVRLCEELTKYXLVTFLGLYVNNQITDIGARYVAQILDE--CRGLKHLKLG 792
Qy   844 LAGCGLTAEDCDLAFGLRANQTLTELDLSFNVLTDAGAXHLCQRLRQPSCK-LQRLQLV 902
Db   793 KNRITSEGGSKCVALAVKNSTSIVDVGMWGNQIGDBGAKAPAEALKDHPS--LTTLSLAFN 850
Qy   903 SCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLR--HPACKLIRLGLDQT 961
Db   851 GISPEGGKSLAQALKQNTILTVIW---------LTKNELNDESAECPAEMLRVNQT 897
Qy   962 TLSDEMRQEL-RALEQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRORLGSERA 1020
Db   898 LRHL 901
Qy  1021 ASHV 1024
```

```
;
; SEQ ID NO 43
;   LENGTH: 953
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-207-359B-43

Query Match            4.6%;  Score 348;  DB 4;  Length 953;
Best Local Similarity  21.6%;  Pred. No. 3.4e-24;
Matches  195;  Conservative 130;  Mismatches 361;  Indels 218;  Gaps 32;

Qy  234  WAAVVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDPNQKFTQLLLLQRP------  287
Db  103  WLSEIGFSP---------------------------SQLIRKTIVNTDPVSRYTQQLRHQLGRDSKFM  144

Qy  288  --HPRSQDPLVKRSWPD------------YVEENRGHLIEIRDLF--GPGLDTQEPRIVILQGAA  336
Db  145  LCYAQKEDLLEETYMDTLMGLVGFNNENLGSLGGLDCLLLDHSTGVLNEHGETVFVFGDA  204

Qy  337  GIGKSTLARQVKEAWGRCQLYGDRFQHVFYPSCRELA------QSKVVSLAELIGKD---GTA  390
Db  205  GVGKSMLLQRLQSLWASGRLTSTA-KFFHFHRCRMFSCFKESDMLSLQDLLFKHCYPEQ  263

Qy  391  TPAPIRQILSR-PERLLFILDGVDE-----PGWVLQEPSSELCLHHSQPQPADALLGSLLGK  446
Db  264  DPEEVFSFLLRFPHTALPTPDGLDELHSDFDLSRVPDS--CCFWEPAHPL-VLLANLLSG  320

Qy  447  TILPEASFLITARTTALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRL  506
Db  321  RLLKGAGKLLTARTGV----EVPRQLLRKKVLLRGFSPSHLRAYARRMFPERTAQEHLLQQ  377

Qy  507  VKSNKELWALCLLVPWVSWLACTCLMQ----------------------------QMKRK  537
Db  378  LDANPNLCSLCGVPLFCWIIPRCFQHFQTVFEGSSSSQLPDCAVTLTDVFLLVTEVHLNRP  437

Qy  538  EKLTLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAABGIWQKKTLFSPDDLRKHGLD  597
Db  438  QPSSLVQRNTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEEVQASKLQ  490

Qy  598  GAIISTFLKMGILQEHP-----IPLSYSPIHLCFQEPFAAMSVLEDE-----------  640
Db  491  -----EGDLQLGFLRALPDVGPEQGQSYEFFHLTLQAFTAFFLVADDKVSTRELLRFFRE  546

Qy  641  ----KGRGKHSNCIIDLEKTLEAYGIHGLFGAS----------TTRFLLGLLSDEG  682
Db  547  WTSPGEATSSSCHSSF------FSFQCLGGRSRLGPDPPRNKDHFQPTNLFVCGLLAKAR  600

Qy  683  EREMENIFHCRLSQGGRNLMQW-----------VPSL------QLLLQPHSLESIHCLY  723
Db  601  QKLLRQLVPKAILRRKRKALWAHLFASLRSYLKSLPRVQSGGPNQVHAMPTFLWMLRCIY  660

Qy  724  ETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTM  783
Db  661  ETQ-----SQKVGRLAARGI------SADYLKIAPCNACSAADCSALSPVLHHFHRQL----  706

Qy  784  VVLFRWVPVTDAYWQILLFSVLKVTRNLKELDLSGNSLSHSAVRSLCKTLRRPCLLETLR  843
Db  707  ------------ALDLDNNNLNDYGVQELQPCPSR-----LTVIR  734

Qy  844  LAGCGLTAEDCKDLAPGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCK-LQRLQLV  902
Db  735  LSVNQITDTGVKVLCEELTKYKIVTFLGLYNNQTTDIGARYVAQILDE--CRGLKHLKLG  792

Qy  903  SCGLTSDCCQDLLASVLSASPSLKRLDLLQQNNLDDVGVRLLCEGLR--HPACKLIRLGLDQT  961
Db  793  KRITSEGGKCVALAVKNSTSIVDVGWMGNQIGDEGAKAFAPEALKDHPS--LTTLSLAFN  850

Qy  962  TLSDEMRQEL-RALEQEKPQLLIFSRRKFSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA  1020
Db  851  GISPEGGKSLAQALKQNTTLTVIW------LITKNELNDESAECFAEMLRVNQT  897

Qy  1021 ASHV  1024
Db  898  LRHL  901
```

US-09-340-620A-43

```
Query Match            4.6%;  Score 348;  DB 4;  Length 953;
Best Local Similarity  21.6%;  Pred. No. 3.4e-24;
Matches  195;  Conservative 130;  Mismatches 361;  Indels 218;  Gaps 32;

Qy  234  WAAVVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDPNQKFTQLLLLQRP------  287
Db  103  WLSEIGFSP---------------------------SQLIRKTIVNTDPVSRYTQQLRHQLGRDSKFM  144

Qy  288  --HPRSQDPLVKRSWPD------------YVEENRGHLIEIRDLF--GPGLDTQEPRIVILQGAA  336
Db  145  LCYAQKEDLLEETYMDTLMGLVGFNNENLGSLGGLDCLLLDHSTGVLNEHGETVFVFGDA  204

Qy  337  GIGKSTLARQVKEAWGRCQLYGDRFQHVFYPSCRELA------QSKVVSLAELIGKD---GTA  390
Db  205  GVGKSMLLQRLQSLWASGRLTSTA-KFFHFHRCRMFSCFKESDMLSLQDLLFKHCYPEQ  263

Qy  391  TPAPIRQILSR-PERLLFILDGVDE-----PGWVLQEPSSELCLHHSQPQPADALLGSLLGK  446
Db  264  DPEEVFSFLLRFPHTALPTPDGLDELHSDFDLSRVPDS--CCFWEPAHPL-VLLANLLSG  320

Qy  447  TILPEASFLITARTTALQNLIPSLEQARWVEVLGFSESSSRKEYFYRYFTDERQAIRAFRL  506
Db  321  RLLKGAGKLLTARTGV----EVPRQLLRKKVLLRGFSPSHLRAYARRMFPERTAQEHLLQQ  377

Qy  507  VKSNKELWALCLLVPWVSWLACTCLMQ----------------------------QMKRK  537
Db  378  LDANPNLCSLCGVPLFCWIIPRCFQHFQTVFEGSSSSQLPDCAVTLTDVFLLVTEVHLNRP  437

Qy  538  EKLTLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAABGIWQKKTLFSPDDLRKHGLD  597
Db  438  QPSSLVQRNTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEEVQASKLQ  490

Qy  598  GAIISTFLKMGILQEHP-----IPLSYSPIHLCFQEPFAAMSVLEDE-----------  640
Db  491  -----EGDLQLGFLRALPDVGPEQGQSYEFFHLTLQAFTAFFLVADDKVSTRELLRFFRE  546

Qy  641  ----KGRGKHSNCIIDLEKTLEAYGIHGLFGAS----------TTRFLLGLLSDEG  682
Db  547  WTSPGEATSSSCHSSF------FSFQCLGGRSRLGPDPPRNKDHFQPTNLFVCGLLAKAR  600

Qy  683  EREMENIFHCRLSQGGRNLMQW-----------VPSL------QLLLQPHSLESIHCLY  723
Db  601  QKLLRQLVPKAILRRKRKALWAHLFASLRSYLKSLPRVQSGGPNQVHAMPTFLWMLRCIY  660

Qy  724  ETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTM  783
```

```
RESULT 14
US-08-519-547A-6
Sequence 6, Application US/0851954
Patent No. 5994082
GENERAL INFORMATION:
APPLICANT:
   TITLE OF INVENTION: Proteins Essential for the Expression of
   TITLE OF INVENTION: Vertebrate MHC Class II Genes, DNA Sequences Encoding Same
   TITLE OF INVENTION: and Pharmaceutical Compositions
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
   ADDRESSEE: FISH & NEAVE
   STREET: 1251 AVENUE OF THE AMERICAS
   CITY: NEW YORK
   STATE: NEW YORK
   COUNTRY: U.S.A.
   ZIP: 10020-1104
COMPUTER READABLE FORM:
   MEDIUM TYPE: Floppy disk
   COMPUTER: IBM PC compatible
   OPERATING SYSTEM: MS-DOS
   SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/519,547A
   FILING DATE: 25-AUG-1995
   CLASSIFICATION: 435
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: EP94113378.7
   FILING DATE: 26-AUG-1994
ATTORNEY/AGENT INFORMATION:
   NAME: HALEY, JAMES F.
   REGISTRATION NUMBER: 27,794
   REFERENCE/DOCKET NUMBER: VOS-11
TELECOMMUNICATION INFORMATION:
   TELEPHONE: 212-596-9000
   TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
   LENGTH: 1130 amino acids
   TYPE: amino acid
   STRANDEDNESS: double
   TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
```

```
Query Match
Best Local Similarity

4.6%;        Score 345;  DB 2;  Length 1130;
23.2%;       Pred. No. 9.3e-24;
```

Matches 256;　Conservative 134;　Mismatches 430;　Indels 282;　Gaps 49;

```
Qy    926  ELDLQQNNLDDVGVRLLCEGLR 947
Db   1076  VMDVQYNKFTAAGAQQLAASLR 1097
```

RESULT 15
US-09-207-359B-47
Sequence 47, Application US/09207359B
Patent No. 6469140
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-112001
CURRENT APPLICATION NUMBER: US/09/207,359B
CURRENT FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 47
LENGTH: 966
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: majority sequence
US-09-207-359B-47

Query Match 4.3%; Score 326.5; DB 4; Length 966;
Best Local Similarity 21.9%; Pred. No. 4.5e-22;
Matches 187; Conservative 113; Mismatches 325; Indels 227; Gaps 31;

```
Qy   271  NEDPNQKFTQLLLLQRP-----------HPRSQDPLVKRSWPD-----YVEENRGHLIEIR 315
Db   122  NTDPVSRYTQQLRHQLGRDSKFVLCYAQKEDLLLEEIVMDTLMGLVGFSNESLSGSLGGLA 181

Qy   316  DLF--GPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQL-YGDRFQHVFYFSCREL 372
Db   182  CLLDHSTGVLNEQGETVFVLGDAGVGKSMLLQRLQSLWASGRLITAGAKF--FFHPRCRMF 239

Qy   373  A---QSKVVSLAELIGKD---GTATPAPIRQILSR--PERLLFLLDGVDEPGWVLQEPSSE 425
Db   240  SCFKESDRLSLQDLLFKHFCYPEQPEEVPAFLLRFPHVALFTFDGLD------ELHSD 292

Qy   426  LCLH-------WSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVE 477
Db   293  LDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGAGKLLTARTGV--EVPRQLLRKKVL 348

Qy   478  VLGFSESSSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPMWSWLACTCLMQ----- 532
Db   349  LRGFSPSHLRAYAPRMFPERAAQDHLLSQLDANPNLCSLCGPVLFCWIIPRCFQHPQAAP 408

Qy   533  --------------QMKREKLILTSKTTTTLC--LHYLAQALQAQPLGP 566
Db   409  BGSSSQLPDCAVTLTDVFLLVTEVHLNRMQPSSLVQRNTRSPAETLHAGRDTLHA----- 463

Qy   567  QLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHP----IPLSYS 620
Db   464  ----LGEVAHRGTDKSLPVFGQEEVQASGLQ----EGDLQLGFLIRALPDVGPGGDGQSYE 515

Qy   621  FIHLCFQEFFAAMSYVLEDEKGRGK---------HSNCIIDLEKTLEAYG 661
Db   516  FFHLTLQAFTAFFLVADDKVGTQELLRFFQEWTSPGGAASSCHSSFL----SFQCLG 570

Qy   662  IHGLFGASTTRFLLGLLSDEGEREMENIFHCRL---SQGRNLMQWVPSL-------- 707
Db   571  GSGRAGED------LFKNDHFQTNLFVCGLLAKAKQKLLRQLVPAAALRRKKALWA 623

Qy   708  ------------------QLLLQPHSLESLHCLYETRNKIFLTQVMAHFEEMGMCV 745
Db   708  ------- 708
```

```
Db   624  HLFASLRGYLKSLPRVQVGGFNQVQAMPTFLWMLRCIYETQ-----SQKYGQLAARGI-- 676
Qy   746  ETDMELLVCTFCIKFSRHVKLLQLIEGRQHRSTWNSPTMVVLFRWVPVTDAYWQILFSVLK 805
Db   677  --SADYLKLAFCNACSADCSALSFVLHHFHKQL------------ 707

Qy   806  VTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKDIAFGLRANQ 865
Db   708  ------ALDLDNNLNDYGVQELQPCFSR----LTVLRLSVNQITDGGVKVLSEELTKYK 757

Qy   866  TLITELDLSFNVLTDRAGAKHLCQRLRQPSCK-------LQRLQLVSCG------LTTSDCCQ 912
Db   758  IVTFLGLYNNQITDVGARYVAQILDE--CKGLTHLSLYNNQITDVGAKLGKNKITSEGGK 815

Qy   913  DLASVLSAPPSLKELDLQQNNLDDVGVRLLCEGLR-HPACKLIRLGLDDQTTLSDEMRQEL 971
Db   816  YVALAVKNSTSIVDVGMWGNQVGDEGAKAFAEALKDHPS--LTTLSLASNGISTEGGKSL 873

Qy   972  RALEQSKPQLLI 983
Db   874  AQALQQNTSLTV 885
```

Search completed: January 29, 2004, 13:48:52
Job time : 31.1576 secs

```
OM protein - protein search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:             January 29, 2004, 13:46:01 ; Search time 249.019 Seconds
                                                (without alignments)
                                                1192.602 Million cell updates/sec

Title:              7534
Perfect score:      7534
Sequence:           1 MAGGAWGRLACYLEFLKKEE.........HLIMELWEKGSKKGLLPLSS 1429

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           789580 seqs, 207824079 residues
Total number of hits satisfying chosen parameters:        789580

Post-processing: Minimum Match 0%
                 Maximum Match 100%
Minimum DB seq length: 0
Maximum DB seq length: 200000000
                 Listing first 45 summaries

Database :       Published Applications AA:*
                 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 2: /cgn2_6/ptodata/1/pubpaa/Pct_NEW_PUB.pep:*
                 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6: /cgn2_6/ptodata/1/pubpaa/PCTU5_PUBCOMB.pep:*
                 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7534	100.0	1429	10	US-09-996-617-2	Sequence 2, Appl
2	7534	100.0	1429	10	US-09-931-071-2	Sequence 2, Appli
3	7534	100.0	1429	12	US-10-028-374-15	Sequence 15, Appl
4	7534	100.0	1429	12	US-10-183-770-15	Sequence 15, Appl
5	7534	100.0	1473	15	US-10-028-392-11	Sequence 11, Appl
6	7502	99.6	1473	10	US-09-388-221-2	Sequence 2, Appl
7	7488	99.4	1429	12	US-10-028-374-3	Sequence 3, Appli
8	7488	99.4	1429	12	US-10-183-770-3	Sequence 3, Appli
9	7364	97.7	1399	10	US-09-388-221-4	Sequence 4, Appli
10	7332	97.3	1443	10	US-09-388-221-6	Sequence 6, Appli
11	6344	84.2	1454	10	US-09-388-221-10	Sequence 10, Appl
12	6174	81.9	1424	10	US-09-388-221-12	Sequence 12, Appl
13	4072	54.0	764	12	US-10-407-866-92	Sequence 92, Appl
14	2171.5	28.8	442	11	US-10-895-298-139	Sequence 139, App
15	1382.5	18.4	1027	12	US-10-407-866-68	Sequence 68, Appl
16	1380	18.3	1061	15	US-10-066-521-18	Sequence 18, Appl
17	1375	18.3	1035	11	US-09-965-621-24	Sequence 24, Appl
18	1375	18.3	1035	10	US-10-407-866-24.	Sequence 24, Appl
19	1364	18.1	1004	12	US-10-108-260A-3161	Sequence 3161, Ap
20	1279	17.0	1034	12	US-10-028-374-18	Sequence 18, Appl
21	1279	17.0	1034	11	US-10-132-967-5	Sequence 5, Appli
22	1279	17.0	1034	12	US-10-183-770-18	Sequence 18, Appl
23	1279	17.0	1034	14	US-10-127-516-5	Sequence 5, Appl
24	1279	17.0	896	15	US-09-965-621-22	Sequence 22, Appl
25	1154	15.3	896	15	US-10-066-521-22	Sequence 22, Appl
26	1037	13.8	994	11	US-09-965-621-16	Sequence 16, Appl
27	1037	13.8	994	14	US-10-407-866-16	Sequence 16, Appl
28	1037	13.8	994	12	US-10-066-521-24	Sequence 24, Appl
29	1008	13.4	1162	15	US-10-066-521-16	Sequence 2718, Ap
30	994	13.1	635	12	US-10-407-866-90	Sequence 90, Appl
31	973.5	12.9	858	9	US-09-848-035-8	Sequence 8, Appli
32	972.5	12.9	858	10	US-09-986-224-8	Sequence 8, Appli
33	972.5	12.9	1143	15	US-10-066-521-6	Sequence 6, Appli
34	969	12.9	1062	12	US-10-239-663-43	Sequence 43, Appl
35	948	12.6	1033	12	US-10-239-663-4	Sequence 4, Appl
36	945.5	12.5	1033	14	US-10-127-516-2	Sequence 2, Appl
37	945.5	12.5	1033	14	US-10-027-629-2	Sequence 2, Appl
38	945.5	12.5	1344	15	US-10-066-521-6	Sequence 6, Appli
39	945.5	12.5	1016	10	US-09-986-224-19	Sequence 19, Appl
40	943.5	12.5	1049	12	US-10-239-663-42	Sequence 42, Appl
41	941.5	12.5	1033	12	US-10-239-663-42	Sequence 42, Appl
42	924.5	12.3	732	12	US-10-407-866-72	Sequence 72, Appl
43	922.5	12.3	952	12	US-10-407-866-70	Sequence 70, Appl
44	913.5	12.1	674	12	US-10-407-866-89	Sequence 89, Appl
45	908.5	12.1	980	9	US-09-848-035-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-09-996-617-2;
Sequence 2, Application US/09996617
Patent No. US2002012819BA1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-340001
CURRENT APPLICATION NUMBER: US/09/996,617
CURRENT FILING DATE:
CURRENT APPLICATION NUMBER: 2001-11-27
PRIOR APPLICATION NUMBER: 09/931,071
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/428,252
PRIOR FILING DATE: 1999-10-27
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
  LENGTH: 1429
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-996-617-2

Query Match        100.0%;   Score 7534;  DB 10;   Length 1429;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1429;  Conservative  0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSFPYSPSEPHLGSPSQPTSTRAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSFPYSPSEPHLGSPSQPTSTRAVLMPW 120

Qy  121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESPSQESPNAPTST 180
```

RESULT 2
US-09-931-071-2

; Sequence 2, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-2

Query Match   100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1429; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 MAGGAWGRLACYLEFLKREELKEFQLLLANKAHSRSSSCGETPAQPEKTSGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKREELKEFQLLLANKAHSRSSSCGETPAQPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSOPTSTAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSOPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESPSQESPNAPTST 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESPSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTELRERERREKSEKGRPPWAAAVGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTELRERERREKSEKGRPPWAAAVGT 240

Qy    241 PPQAHTSLQPHHIPWEPSVRESLCSTWPWKNEDPNQKFTQLILLQRPHPRSQDPLVKRSW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHIPWEPSVRESLCSTWPWKNEDPNQKFTQLILLQRPHPRSQDPLVKRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPOPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPOPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
```

; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens

US-10-028-374-15

Query Match   100.0%;   Score 7534;   DB 12;   Length 1429;
Best Local Similarity 100.0%;   Pred. No. 0;   Mismatches 0;   Indels 0;   Gaps 0;
Matches 1429;   Conservative 0;

Qy   1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy   121 IHELPAGCTCQGSERVLRQLPDTSGRRWREISASILYQALPSSPDHESPSQOESPNAPTST 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IHELPAGCTCQGSERVLRQLPDTSGRRWREISASILYQALPSSPDHESPSQOESPNAPTST 180

Qy   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy   241 PPQAHTSLQPHHHWPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 PPQAHTSLQPHHHWPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy   301 PDYEENRGHLIEIRDLFGPGLDTQEPRIVILQGGHVDTSLFQMAHFKEEGMLLBKPARV 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 PDYEENRGHLIEIRDLFGPGLDTQEPRIVILQGGHVDTSLFQMAHFKEEGMLLBKPARV 360

Qy   361 FQHVFYFSCREBLAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 FQHVFYFSCREBLAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG 480

Qy   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy   661 GIHGLFGASTTRFLLGLLSDBGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GIHGLFGASTTRFLLGLLSDBGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy   721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGROHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGROHRSTWS 780

Qy   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy   661 GIHGLFGASTTRFLLGLLSDBGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GIHGLFGASTTRFLLGLLSDBGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy   721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGROHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGROHRSTWS 780

Qy   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy   841 TLRLAGCGLTAEDCKDLAFGLRANQTLTFELDLSFNVLIDAGAKHLCQRLRQPSCKLQRLQ 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TLRLAGCGLTAEDCKDLAFGLRANQTLTFELDLSFNVLIDAGAKHLCQRLRQPSCKLQRLQ 900

Qy   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQONNLDDVGVRLLCBGLRHPACKLIRLGLDQ 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 LVSCGLTSDCCQDLASVLSASPSLKELDLQONNLDDVGVRLLCBGLRHPACKLIRLGLDQ 960

Qy   961 TTLSDBMRQELRALBQEKPQLLIFSRRKPSVMPTPEGLDTGEMSNSTSSLKRQRLGSBRA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TTLSDBMRQELRALBQEKPQLLIFSRRKPSVMPTPEGLDTGEMSNSTSSLKRQRLGSBRA 1020

Qy   1021 ASHVAQANLKLLDVSKIFPIAEIABESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ASHVAQANLKLLDVSKIFPIAEIABESSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy   1081 GPTGPVATEVVDKEKNLVRVHFPVAGSYRWPNTGLCFVMREAVTVEIEPCVWDQFLGEIN 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GPTGPVATEVVDKEKNLVRVHFPVAGSYRWPNTGLCFVMREAVTVEIEPCVWDQFLGEIN 1140

Qy   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLBKPARV 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLBKPARV 1200

Qy   1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPBEEVTFHLYLIPSDCSIR 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPBEEVTFHLYLIPSDCSIR 1260

Qy   1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320

Qy   1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLABNTRPS 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLABNTRPS 1380

Qy   1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429

```
RESULT 4
US-10-183-770-15
Sequence 15, Application US/10183770
Publication No. US20030180812A1
GENERAL INFORMATION:
  APPLICANT: Bristol-Myers Squibb Company
  TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
  TITLE OF INVENTION:   PREDOMINATELY IN BONE MARROW, HLRRBM1
  FILE REFERENCE: D0067A.CIP
  CURRENT APPLICATION NUMBER: US/10/183,770
  CURRENT FILING DATE: 2002-06-27
  PRIOR APPLICATION NUMBER: US 60/257,773
  PRIOR FILING DATE: 2000-12-22
  PRIOR APPLICATION NUMBER: US 10/028,374
  PRIOR FILING DATE: 2001-12-20
  NUMBER OF SEQ ID NOS: 33
  SOFTWARE: Patentin version 3.1
  SEQ ID NO 15
  LENGTH: 1429
  TYPE: PRT
  ORGANISM: homo sapiens
US-10-183-770-15

Query Match            100.0%;    Score 7534;   DB 12;   Length 1429;
Best Local Similarity  100.0%;    Pred. No. 0;
Matches 1429;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ   60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFYPSPSEPHLGSPSPSQOPTSTAVLMPW  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFYPSPSEPHLGSPSPSQOPTSTAVLMPW  120
```

(sequence alignment continues with matching query/database rows from positions 121 through 1429)

```
Db  1141  POHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMLLEKPARV  1200

Qy  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy  1261  KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP  1320

Qy  1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380

Qy  1381  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
```

RESULT 5
US-10-028-392-11
; Sequence 11, Application US/10028392
; Publication No. US20030087340A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN NERVOUS SYSTEM TISSUES, HLRRNS1
; FILE REFERENCE: D0085.np
; CURRENT APPLICATION NUMBER: US/10/028,392
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/259,479
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US 60/260,616
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 11
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-392-11

Query Match   100.0%;   Score 7534; DB 15; Length 1429;
Best Local Similarity 100.0%;   Pred. No. 0; Mismatches 0;
Matches 1429; Conservative 0; Indels 0; Gaps 0;

```
Qy   1  MAGGAWGRIACYLBFLKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1  MAGGAWGRLACYLEFLKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ   60

Qy  61  YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61  YGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW  120

Qy 121  IHELPAGCTQGSERRVLRQLPDTSGIYYTEIRERERKSEKGRPPWAAVGT  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121  IHELPAGCTQGSERRVLRQLPDTSGIYYTEIRERERKSEKGRPPWAAVGT  180

Qy 181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVGT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVGT  240

Qy 241  PPQAHTSLQPHHHWEPSVRESLCGSTWPWKNEDFNQKFTQLLLQRPHPRSQDPLVKRSW  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241  PPQAHTSLQPHHHWEPSVRESLCGSTWPWKNEDFNQKFTQLLLQRPHPRSQDPLVKRSW  300

Qy 301  PDYEENRGHLIEIRDLFGPGLDTQEPRIVILQGHAGIGKSTLARQVKEAWGRGQLYGDR  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301  PDYEENRGHLIEIRDLFGPGLDTQEPRIVILQGHAGIGKSTLARQVKEAWGRGQLYGDR  360

Qy 361  FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361  FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ  420
```

```
Qy 421  EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421  EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWEVLG  480

Qy 481  FSESSRKEYFYRYFPDERQAIRAFRLVKSNKELWALCLVPWNVSWLACTCLMQQMRKEKL  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481  PSESSRKEYPYRYFTDERQAIRAFRLVKSNKELWALCLVPWNVSWLACTCLMQQMRKEKL  540

Qy 541  TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541  TLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI  600

Qy 601  ISTPLKMGILQEHPIPLSYSFIHLCFQEEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601  ISTPLKMGILQEHPIPLSYSFIHLCFQEEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660

Qy 661  GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLBSLH  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661  GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLBSLH  720

Qy 721  CLYEYRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721  CLYEYRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS  780

Qy 781  PTMWVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781  PTMWVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE  840

Qy 841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ  900

Qy 901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ  960

Qy 961  TTLSDEMRQEIRALRQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA  1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961  TTLSDEMRQEIRALRQEKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA  1020

Qy 1021 ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW  1080

Qy 1081 GPTGPVATEVVDKERQNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GPTGPVATEVVDKERQNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN  1140

Qy 1141 POHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMLLEKPARV  1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 POHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMLLEKPARV  1200

Qy 1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR  1260

Qy 1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP  1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 KELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPP  1320

Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS  1380

Qy 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS  1429
```

RESULT 6
US-09-388-221-2
; Sequence 2, Application US/09388221A

```
; Publication No. US20020192643A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, John C.
;  TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
;  FILE REFERENCE: P-LJ 3650
;  CURRENT APPLICATION NUMBER: US/09/388,221A
;  CURRENT FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
;  LENGTH: 1473
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-388-221-2

  Query Match             99.6%;  Score 7502;  DB 10;  Length 1473;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches 1429;  Conservative    0;  Mismatches    0;  Indels   44;  Gaps    1;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETPAQEPEKTSGMEVASYLVAQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETPAQEPEKTSGMEVASYLVAQ 60

Qy     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESSPSQESPNAPTST 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESSPSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHPHPWEPSVRESLCGSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDPLVRSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHPHPWEPSVRESLCGSTWPWKQNEDFNQKFTQLLLLQRPHPRSQDPLVRSW 300

Qy    301 PDVYEENRGHLIEIRDLFGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDVYEENRGHLIEIRDLFGPGLDTQEEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALCQNLIPSLEQARWVEVLG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALCQNLIPSLEQARWVEVLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLYPWVSWLACTCLMQQMRKEKL 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLYPWVSWLACTCLMQQMRKEKL 540

Qy    541 TLISKTTTTLCLHYLAQALLQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLISKTTTTLCLHYLAQALLQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLFGASTTRFLILGLLSDEGEREMENIFHCRLSQGRNLMQWTPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLFGASTTRFLILGLLSDEGEREMENIFHCRLSQGRNLMQWTPSLQLLLQPHSLESLH 720

Qy    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLLVCTFCIKFSRHVKKLQLLIEGRQHRSTWS 780

Qy    781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
```

```
Db    781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ 960

Qy    961 TTLISDEMRQELRALEQEKPQLLIFSRRKPSVMTTPEGLDTGEMSNSTSSLKRQRLGSERA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTLISDEMRQELRALEQEKPQLLIFSRRKPSVMTTPEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy   1021 ASHVAQANLKLLDVSKIFPIAEIABESSSPEVVPYELLCYPSPASQGDLHTKPLGTDDDFW 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ASHVAQANLKLLDVSKIFPIAEIABESSSPEVVPYELLCYPSPASQGDLHTKPLGTDDDFW 1080

Qy   1081 GPTGPVATEVVDKEKNLYRVHFPVYAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GPTGPVATEVVDKEKNLYRVHFPVYAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140

Qy   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDITSLFQWAHFKEEGMLLEKPARV 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDITSLFQWAHFKEEGMLLEKPARV 1200

Qy   1201 ELHHIVLENPSFSPLGVLKMIHNALRFIPVTSVVLLYHRVVHPEEVTFHLYLIPSDCSIR 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ELHHIVLENPSFSPLGVLKMIHNALRFIPVTSVVLLYHRVVHPEEVTFHLYLIPSDCSIR 1260

Qy   1261 --------------------------------------------------KELELCYRSPGEDQLF 1276
                                                              ||||||||||||||||
Db   1261 KAIDDLEMKFQFVRIHKPPPLTPLYMGCRYTVSGSGSGMLEILPKELELCYRSPGEDQLF 1320

Qy   1277 SEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 SEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1380

Qy   1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKC 1396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKC 1440

Qy   1397 KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
          ||||||||||||||||||||||||||||||||
Db   1441 KDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1473
```

```
RESULT 7
US-10-028-374-3
; Sequence 3, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
;  TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRREM1
;  FILE REFERENCE: D0067NP
;  CURRENT APPLICATION NUMBER: US/10/028,374
;  CURRENT FILING DATE: 2001-12-20
;  PRIOR APPLICATION NUMBER: US 60/257,773
;  PRIOR FILING DATE: 2000-12-22
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
;  LENGTH: 1429
;  TYPE: PRT
;  ORGANISM: homo sapiens
US-10-028-374-3

  Query Match             99.4%;  Score 7488;  DB 12;  Length 1429;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1420;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETPAQEPEKTSGMEVASYLVAQ 60
```

LEFT COLUMN — alignment (query vs database), positions 1–1140:

```
Db    1  MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQEPEKTSGMEVASYLVAQ   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1  MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQEPEKTSGMEVASYLVAQ   60

Db   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSSEPHLGSPSQPTSTAVLMPW  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSSEPHLGSPSQPTSTAVLMPW  120

Db  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180

Db  181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT  240

Db  241  PPQAHSSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300
         ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241  PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300

Db  301  PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301  PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR  360

Db  361  FQHVFYPSCRBLAQSKVVSLAELIGHDGTATPAPIRQILSRPERLLFILDGVDDEPGWVLQ  420
         |||| :| :|||||||||||||| :||||||| :||||||||||||||||||||||: |||
Qy  361  FQHVYFYSCRBLAQSKVVSLAELICKDGTATPAITPAPIRQILSRPERLLFILDGVDDPGWVLQ  420

Db  421  EPSSELCLHWSQPQPADALLGSLLGKTILPEASFIITAARTTALQNLIPSLEQARWEVLG  480
         ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Qy  421  EPSSELCLHWSQPOPADALLGSLLGKTILPEASFIITAARTTALQNLIPSLEQARWEVLG  480

Db  481  FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481  FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL  540

Db  541  TLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDLRKHGLDGAI  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541  TLTSKTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDLRKHGLDGAI  600

Db  601  ISTFLKMGILQEHPIPLSYSFIHLCFQEBPFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601  ISTFLKMGILQEHPIPLSYSFIHLCFQEBPFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY  660

Db  661  GIHGLFGASTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLILLQPHSLBSLH  720
         ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Qy  661  GIHGLFGASTRFLIGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLILLQPHSLBSLH  720

Db  721  CLYETRNKFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKLQLILEGRQHRSTWS  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721  CLYETRNKFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKLQLILEGRQHRSTWS  780

Db  781  PSMVVLFRNWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE  840
         | :||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781  PTMVVLFRNWPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE  840

Db  841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLMDPAGAKHLCQRLRQPSCKLQRLQ  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841  TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLMDPAGAKHLCQRLRQPSCKLQRLQ  900

Db  901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901  LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRIGLDQ  960

Db  961  TTLLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPEGLDTGEMSNSTSSLKRORLGSERA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961  TTLLSDEMRQELRALEQEKPQLLIFSRRKPSVMTPEGLDTGEMSNSTSSLKRORLGSERA 1020

Db 1021  ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVELLCVPSPASQGGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021  ASHVAQANLKLLDVSKIFPIAEIAEBSSPEVVPVELLCVPSPASQGGDLHTKPLGTDDDFW 1080

Qy 1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVVREAVTVEIEFCVWDQFLGEIN 1140
```

RIGHT COLUMN — continuation of alignment (positions 1081–1429):

```
Db 1081  GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVVREAVTVEIEFCVWDQFLGEIN 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFPQMAHFKEEGMLLEKPARV 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFPQMAHFKEEGMLLEKPARV 1200

Qy 1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLIYHVHPEEVTFHLYLIPSDCSIR 1260
         |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db 1201  ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLIYHRLHPEEVTFHLYLIPSDCSIR 1260

Qy 1261  KELELCYRSPGGEDQLFSEFYVGHLGSGIRLQVKDKDDETLLWEALVKPGDLMPATTLIPP 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  KELELCYRSPGGEDQLFSEFYVGHLGSGIRLQVKDKDDETLLWEALVKPGDLMPATTLIPP 1320

Qy 1321  ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
         |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  ACIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy 1381  QMRKLFSLSQSWDRKCKDGLIYQALKETHPHLLIMELWEKGSKKGLLFLSS 1429
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  QMRKLFSLSQSWDRKCKDGLIYQALKETHPHLLIMELWEKGSKKGLLFLSS 1429
```

RESULT 8
US-10-183-770-3

```
; Sequence 3, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRBM1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-3

Query Match         99.4%;    Score 7488;    DB 12;    Length 1429;
Best Local Similarity    99.4%;
Matches 1420;   Conservative    5;   Mismatches    4;    Indels    0;    Gaps    0;
```

RESULT 8 — alignment (positions 1–360):

```
Qy    1  MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQEPEKTSGMEVASYLVAQ   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQEPEKTSGMEVASYLVAQ   60

Qy   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSSEPHLGSPSQPTSTAVLMPW  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSSEPHLGSPSQPTSTAVLMPW  120

Qy  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST  180

Qy  181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT  240

Qy  241  PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300
         ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  PPQAHSSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW  300

Qy  301  PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR  360
```

Qy 361 FQHVFYPSCRELAQSKVVSLAELIGKDPGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 FQHVFYPSCRELAQSKVVSLAELIGKDPGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy 421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540

Qy 541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy 601 ISTFLRMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 ISTFLRMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy 661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy 721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy 781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840
         |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 PSMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLE 840

Qy 841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLMDAGAKHLCQRLRQPSCKLQRLQ 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLMDAGAKHLCQRLRQPSCKLQRLQ 900

Qy 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy 961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWMTPIEGLDTGEMSNSTSSLKRQRLGSSRA 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWMTPIEGLDTGEMSNSTSSLKRQRLGSSRA 1020

Qy 1021 ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080

Qy 1081 GPTGGPVATEVUDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
         ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db 1081 GPTGGPVATEVUDKEKNLYRVHFPVAGSYRWPNTGLCFVVREAVTVEIEFCVWDQFLGEIN 1140

Qy 1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQVAHFKEEGMLLEKPARV 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQVAHFKEEGMLLEKPARV 1200

Qy 1201 ELHHIVLENPSPSPLGVLLKMIHNALRFIPVTSVVLLYHRLHPEEVTFHLYLIPSDCSIR 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 ELHHIVLENPSPSPLGVLLKMIHNALRFIPVTSVVLLYHRLHPEEVTFHLYLIPSDCSIR 1260

Qy 1261 KELELCYRSPGEDQLPSEFYVGHLGSGIRLQVKDCKDETLVWEALVKPGDLMPATTLLPP 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 KELELCYRSPGEDQLPSEFYVGHLGSGIRLQVKDCKDETLVWEALVKPGDLMPATTLLPP 1320

Qy 1321 ARIAVPSPLDAPQLLHFVDVQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 ACIAVPSPLDAPQLLHFVDVQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380

Qy 1381 QMRKLFSLSQSWDRKCCDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 QMRKLFSLSQSWDRKCCDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429

Query Match   97.7%; Score 7364; DB 10; Length 1399;
Best Local Similarity 97.9%;
Matches 1399; Conservative 0; Mismatches 0; Indels 30; Gaps 1;

Qy   1 MAGGAWGRLACYLEFLKKEELKKPQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MAGGAWGRLACYLEFLKKEELKKPQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQPTSTAVLMPW 120

Qy 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTST 180

Qy 181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERREKSEKGRPHWAAVGT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERREKSEKGRPHWAAVGT 240

Qy 241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy 301 PDIVEENRGHLIEIRDLFGPGLDTQEPRIVILQGGAAGIGKSTLARQVKEAWGRQLYGDR 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 PDIVEENRGHLIEIRDLFGPGLDTQEPRIVILQGGAAGIGKSTLARQVKEAWGRQLYGDR 360

Qy 361 FQHVFYFSCRELAQSKVVSLAELIGKDPGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 FQHVFYFSCRELAQSKVVSLAELIGKDPGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy 421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480

Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540

Qy 541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy 601 ISTFLRMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 ISTFLRMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy 661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy 721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

RESULT 10
US-09-388-221-6
Sequence 6, Application US/09388221A
Publication No. US20020126643A1
GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-IJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match   97.3%;
Best Local Similarity   95.0%;
Matches 1399; Conservative 0; Mismatches 0; Indels 74; Gaps 2;
Score 7332; DB 10; Length 1443;

```
Qy  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy  781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLG-- 957

Qy  961 TTLSDEMRQELRALEQBKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
                                ||||||||||||||||||||||||||||||||||
Db  958 ------------------------KPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 990

Qy 1021 ASHVAQANLKLLDVSKIFPIAEIAEESSEPVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  991 ASHVAQANLKLLDVSKIFPIAEIAEESSEPVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1050

Qy 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1051 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1110

Qy 1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1111 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1170

Qy 1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSTR 1260
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1171 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSTR 1230

Qy 1261 KELELCYRSPGEDQLFSEPYVGHLGSGIRLQVYKDKKDDETLVWEALVKPGDLMPATTLIPP 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1231 KELELCYRSPGEDQLFSEPYVGHLGSGIRLQVYKDKKDDETLVWEALVKPGDLMPATTLIPP 1290

Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1291 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPS 1350

Qy 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1429
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1351 QMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS 1399
```

```
Qy   61 YGEQQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQQTTSTAVLMPW 120
        ||||| ||||||||||||||||||||||||||||| |||||||||||||||| |||| |||
Db   61 YGEQQAWDLALHTWEQMGLRSLCAQAQEGAGHSFSFPYSPSEPHLGSPSQBSPNAPTST 120

Qy  121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDHESPSQBSPNAPTST 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSDHESPSQBSPNAPTST 180

Qy  181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERKSEKGRPPWAAVVGT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy  241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy  301 PDYYEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PDYYEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy  361 FQHVFYFSCRELAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FQHVFYFSCRELAQSKVVSLAELIGKDGTRATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy  421 EPSSELCLHWSQPQPADALLGSSLLGKNTILPEASFLITAARTTALQNLIPSLEQARWEVLG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 EPSSELCLHWSQPQPADALLGSSLLGKNTILPEASFLITAARTTALQNLIPSLEQARWEVLG 480

Qy  481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELVPWVSWLACTCLMQQMRKEKL 540
        |||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db  481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy  541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKGHGLDGA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKGHGLDGA 600

Qy  601 ISTFLKMGIIQEHPIIPLSYSFIHLCFQEFFAAMSYVLEDKGRGKHSNCIIDLEKTLEAY 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ISTFLKMGIIQEHPIIPLSYSFIHLCFQEFFAAMSYVLEDKGRGKHSNCIIDLEKTLEAY 660

Qy  661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy  781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 PTMVVLFRWVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLG-- 957

Qy  961 TTLSDEMRQELRALEQBKPQLLIFSRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
                                ||||||||||||||||||||||||||||||||||
Db  958 ------------------------KPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERA 990

Qy 1021 ASHVAQANLKLLDVSKIFPIABIAEESSEPVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  991 ASHVAQANLKLLDVSKIFPIABIAEESSEPVVPVELLCVPSPASQGDLHTKPLGTDDDFW 1050

Qy 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1051 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVWREAVTVEIEFCVWDQFLGEIN 1110

Qy 1141 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
```

```
Qy    1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAGGAWGRLACYLEFLKKEELKEFQLLIANKAHSRSSSGETPAQPEKTSGMEVASYLVAQ 60
```

```
RESULT 11
US-09-388-221-10
Sequence 10, Application US/09388221A
Publication No. US20020192643A1
 GENERAL INFORMATION:
 APPLICANT:
  APPLICANT: Reed, John C.
 TITLE OF INVENTION:
  FILE REFERENCE: P-LJ 3650
  CURRENT APPLICATION NUMBER: US2002019264 3A1e1 Card Proteins Involved in Cell Death Regul
  CURRENT APPLICATION NUMBER: US/09/388,221A
  CURRENT FILING DATE: 1999-09-01
 NUMBER OF SEQ ID NOS: 18
 SEQ ID NO 10
  LENGTH: 1454
  TYPE: PRT
  ORGANISM: Artificial Sequence
 FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-388-221-10

Query Match             84.2%;
Best Local Similarity   84.2%;  Pred. No. 0;
Matches 1229;  Conservative 53;  Mismatches 122;  Indels 56;  Gaps  7;

Qy          1 MAGGAWGRLACYLEFLKEELKEFQLLLANKAHSRSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGGAWGRLACYLEFLKEELKEFQLLLANKAHSRSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy         61 YGEQRAWDLALHTWEQMCLRSLCAQAQEGAGHSPSFPYYSPSEPHLGSPSQPTSTAVLMPW 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YGEQRAWDLALHTWEQMCLRSLCAQAQEGAGHSPSFPYYSPSEPHLGSPSQPTSTAVLMPW 120

Qy        121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQRSPNAPTST 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IHELPAGCTCQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQBSPNAPTST 180

Qy        181 AVLGSWGSPPQPSLAPRRQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AVLGSWGSPPQPSLAPRRQEAPGTQWPLDETSGIYYTEIRERERKSEKGRPPWAAVVGT 240

Qy        241 PPQAHTSLQPHHPWEPSVRESLCSTWPKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPQAHTSLQPHHPWEPSVRESLCSTWPKNEDPNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy        301 PDYVEENRNGHLIEIRDLFGPGLDTQEPRIVILQQAAGIGKSTLARQYKEAWGRGQLYGDR 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PDYVEENRNGHLIEIRDLFGPGLDTQEPRIVILQQAAGIGKSTLARQYKEAWGRGQLYGDR 360

Qy        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
```

```
Qy        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy        421 EPSSELCLHWSQPOPADALLGSLLGKTIIPEASFLITARTTALQNLIFSLEQARWVEVLG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EPSSELCLHWSQPOPADALLGSLLGKTIIPEASFLITARTTALQNLIFSLEQARWVEVLG 480

Qy        481 FSESSRKEYFYRYIFTDERQAIRAPRLVXSNKELMWALCLVPWVSWLACTCLMQQMRKEKL 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FSESSRKEYFYRYIFTDERQAIRAPRLVXSNKELMWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy        541 TLISKTTTTLCLHYLAQALQAQPLGPQLRDICSLAAEGIWQKKTLFSPDDLRKGILDGAI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TLISKTTTTLCLHYLAQALQAQPLGPQLRDICSLAAEGIWQKKTLFSPDDLRKGILDGAI 600

Qy        601 ISTPLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ISTPLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy        661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy        721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy        781 PTMVVLFRNVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PTMVVLFRNVPVTDAYWQILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLE 840

Qy        841 TLRLAGCGLTAEDCKDLAPGLRANQTLTELDLSFNVLIDAGAKHLCQRLRQPSCKLQRLQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TLRLAGCGLTAEDCKDLAPGLRANQTLTELDLSFNVLIDAGAKHLCQRLRQPSCKLQRLQ 900

Qy        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LVSCGLTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960

Qy        961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTPTEGLDTGEMSNSTSSLKRQRLGSERA 1020

Qy       1021 ASHVAQANLKLLDVSKIFPIAEIAEBESSPEVVPVEBLLCVPSPASQGDLHTKPLGTDDDFW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       1021 ASHVAQANLKLLDVSKIFPIAEIAEBESSPEVVPVEBLLCVPSPASQGDLHTKPLGTDDDFL 1080

Qy       1081 GPFGNVDVELIDKSTNRYSVWFPTPAGWVLWSATCLGFLVRDEVTVTIAFGSWSQHLADL 1140
              :||::||:|:|||||||||||||:||:::|||||||||||:||||||||:||||||||:||::
Db       1081 GPEGNVDVELIDKSTNRYSVWFPTPAGWVLWSATCLGFLVRDEVTVTIAFGSWSQHLADL 1140

Qy       1141 QHHBQWLVGGPLPDVTAFPEEAVAEIHLPHPISLQ-GEVDVSWFLVAHFKNEGMVLEHPA 1199
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db       1141 QHHBQWLVGGPLPDVTAFPEEAVAEIHLPHPISLQ-GEVDVSWFLVAHFKNEGMVLEHPA 1199

Qy       1199 RVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVRVHPEEVTFHLYLIPSDCS 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 RVEPFYAVLESPSFSLMGILLRIASGTRLSIPITSNTLIYYHPEDIKFHLYLVPSDAL 1259

Qy       1259 IR------------------------------KELELCYRSPGEDLQF 1276
                                              ||||||||||||||||
Db       1260 LTKAIDDEEDRFHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF 1319

Qy       1277 SEFYVGHLGSGIRLQVRDKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 SKIYAGQMKEPIQLEITEKRHGTLVWDTEVKPVDLQLVAAASAPP-------PFSGAA--- 1369

Qy       1337 FVDQYREQLIARVTSVEVVLDKLH-GQVLSQBOYERVLAENTRPSQMRKLFSLSQSWDRK 1395
              :|||||||||||||||||||||:| ||||||||||||||||||||||||||||||||||
Db       1370 FVKENHRQLQARMGDLKGVLDDLQDNEVLITENEKELVEQEKTRQSKNEAILSMVEKKGDL 1429

Qy       1396 CKDGLYQALKETHPHLIMEL 1415
              ||:|:|||||||||||:|||:|
Db       1430 ...
```

```
Qy       1111 PQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1170
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1115 ...

Qy       1201 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVRVHPEEVTFHLYLIPSDCSIR 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1171 ELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIR 1230

Qy       1261 --------------KELELCYRSPGEDLQF 1276
                            ||||||||||||||||
Db       1231 KAIDDLEMKPQFVRIHKPPLPLYMGCRYVTSGSGSGSGSGMLEILPKELELCYRSPGEDLF 1290

Qy       1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQBOYERVLAENTRPSQMRKLFSLSQSWDRKC 1396
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 SEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1350

Qy       1277 SEFYVGHLGSGIRLQVKDKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 FVDQYREQLIARVTSVEVVLDKLHGQVLSQBOYERVLAENTRPSQMRKLFSLSQSWDRKC 1410

Qy       1397 KDGLYQALKETHPHLIMELWEKGSSKKGLLPLSS 1429
              ||||||||||||||||||||||||||||||||||
Db       1411 KDGLYQALKETHPHLIMELWEKGSSKKGLLPLSS 1443
```

```
RESULT 12
US-09-388-221-12
; Sequence 12, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, John C.
;  TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
;  FILE REFERENCE: P-LJ 3650
;  CURRENT APPLICATION NUMBER: US/09/388,221A
;  CURRENT FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1424
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-388-221-12

  Query Match             81.9%;  Score 6174;  DB 10;  Length 1424;
  Best Local Similarity   82.1%;  Pred. No.0;
  Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

Qy      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETPAQEBKTSGMEVASYLVAQ 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSSGETPAQEBKTSGMEVASYLVAQ 60

Qy     61 YGEQORAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQOPTSTAVLMPW 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 YGEQORAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSQOPTSTAVLMPW 120

Qy    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESSSQESPNAPTST 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSSPDHESSSQESPNAPTST 180

Qy    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERERKSEKRPPWAAVVGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERERERKSEKRPPWAAVVGT 240

Qy    241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 PPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLLLLQRPHPRSQDPLVKRSW 300

Qy    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGRAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWVBLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITARTTALQNIIPSLEQARWVBLG 480

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy    541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEBGIWQKKTLFSPDDLRKHGLDGAI 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAEBGIWQKKTLFSPDDLRKHGLDGAI 600

Qy    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660

Qy    661 GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GIHGLPGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLH 720

Qy    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLIVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CLYETRNKTFLTQVMAHFEEMGMCVETDMELLIVCTFCIKFSRHVKKLQLIEGRQHRSTWS 780

Qy    781 PTMVVLFRWVPVTDAYWQILFSVLKYTRNLKELDLGSNSLSHSAVKSLCKTLRRPRCLLE 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 PTMVVLFRWVPVTDAYWQILFSVLKYTRNLKELDLGSNSLSHSAVKSLCKTLRRPRCLLE 840

Qy    841 TLRLAGCGLTABDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TLRLAGCGLTABDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQ 900

Qy    901 LVSCGLITSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 LVSCGLITSDCCQDLLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLG--- 957

Qy    961 TTLSDEMRQELRALEQEKPQLLIFSRRKPSVWTFPTEGLDTGEMSNSTSSLKRQRLGSSERA 1020
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    958 ---------------KPSVWTFTEGLDTGEMSNSTSSLKRQRLGSSERA 990
```

```
RESULT 13
US-10-407-866-92
; Sequence 92, Application US/10407866
; Publication No. US2004000 2599A1
; GENERAL INFORMATION:
;  APPLICANT: Godzik, Adam
;  APPLICANT: Reed, John C.
;  TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
;  TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
;  FILE REFERENCE: 66654-10(LJ 5755)
;  CURRENT APPLICATION NUMBER: US/10/407,866
;  CURRENT FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: US 60/370,538
;  PRIOR FILING DATE: 2002-04-04
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 764

Qy    991 ASHVAQANLKLLDVSKIFPIABIAEEBSSPEVVPVELLCVPSPSPASQGDLHTKPLGTDDFW 1050
              |||||||||||||||||| ||||||||||||  ||| ||||||| |||||||||||||
Db    991 ASHVAQANLKLLDVSKIFPIABIAEEBSSPEVVPVELLCVPSPSPASQGDLHTKPLGTDDFL 1050

Qy   1051 GPEGNVDVELIDKSTNRYSVWFPTAGMYLWSATGLGPLVRDEVTVTIAFGSWSQHLALDL 1110
                   |||  ||   ||| | | |  | ||  || |  | |  |: ||
Db   1051 NPQHSMWVAGPLLDDIKAEP-GAVEAVHLPHFVALQGGHVDTISLFQMAHFKEBGMLLEKPA 1110

Qy   1111 QHHEQWLVGGPLFDVTAEPEEAVAEIHLPHFISLQ--GEVDVSWFLVAHFKNEGMVLEHPA 1169
               :|:   ||||||||||  ||: |||||||:|    ||||| || ||| |:|:||| |
Db   1111 QHHEQWLVGGPLFDVTAEPEEAVAEIHLPHFISLQ--GEVDVSWFLVAHFKNEGMVLEHPA 1169

Qy   1170 RVEPFYAVLESPSFSLMGILLRIASGTRLSIPITSNTLIYHPHPEDIKFHLYLVPSDAL 1229
              |:|| |:||| |  :| | |||  |:|||| | |:|  |  ||| |  |||| || |
Db   1170 RVEPFYAVLESPSFSLMGILLRIASGTRLSIPITSNTLIYHPHPEDIKFHLYLVPSDAL 1229

Qy   1230 IR------------------------------------------KELELCYRSPGEDQLF 1276
                                                          :|||:|| |||| |
Db   1230 LTKAIDDEEDRFHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF 1289

Qy   1277 SBFYVGHLGSGIRLQVKDKDDETLLWBEALVKPGDLMPATLLIPPARIAVPSPLDAPQLLH 1336
               ||| ||  |  ||  |  ||:  ||| |:||||  ||:  |||| ||| | ||| |
Db   1290 SKFYACQMKBPIQLEITEKRGHGTLVWDTEVKPVDLQIVAASAPP-------PFSGAA--- 1339

Qy   1337 FVDQYRBQLIARVTSVEVVLLDKLH-GQVLSQBQYERVLAENTRPSQMRKLFSLSQSWDRK 1395
               ::|  |  ||||| |  ||||    ||||| |      :|| |  ||||| ||||||
Db   1340 FVKENHRQLQARMGDLKGVLDDLQDNEVLITENEKELVEQEKTRQSKNEALLSMVEKKGDL 1399

Qy   1396 CKDGLYQALKETHPHLIMEL 1415
              |:|:|    |::|  |:|
Db   1400 ALDVLFRSISERDPYLVSYL 1419
```

```
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-407-866-92

Query Match             54.0%;  Score 4072;  DB 12;  Length 764;
Best Local Similarity   100.0%;  Pred. No. 1.3e-314;
Matches  764;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 MAGGAWGRIACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MAGGAWGRIACYLEFLKKEELKEFQLLLANKAHSRSSSGGETPAQPEKTSGMEVASYLVAQ 60

Qy   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSGPTSTAVLMPW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 YGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSEPHLGSPSGPTSTAVLMPW 120

Qy  121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESQPNAPTST 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQESQPNAPTST 180

Qy  181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSEKGRPWAAAVGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSEKGRPWAAAVGT 240

Qy  241 PPQAHTSLCQPHHHHWEPSVRESLCSTWPWKNEDNQKFTQLLLLQRPHPRSQDDLVKRSW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 PPQAHTSLCQPHHHHWEPSVRESLCSTWPWKNEDNQKFTQLLLLQRPHPRSQDDLVKRSW 300

Qy  301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

Qy  361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILLDGVDEPGWVLQ 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 FQHVFYFSCRELAQSKVVSLAELIGKDGTATPAPIRQILSRPERLLFILLDGVDEPGWVLQ 420

Qy  421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITAATTALQNLIPSLEQARWVEVLG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 EPSSELCLHWSQPQPADALLGSLLGKTILPEASFLITAATTALQNLIPSLEQARWVEVLG 480

Qy  481 FSESSRKEYFYRYRYFTDERQAIRAFRLIVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 FSESSRKEYFYRYRYFTDERQAIRAFRLIVKSNKELWALCLVPWVSWLACTCLMQQMRKEKL 540

Qy  541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TLTSKTTTTLCLHYLAQALQAQPLGPQLRDLCSLAAAEGIWQKKTLFSPDDLRKHGLDGAI 600

Qy  601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGRGHSNCLIIDLEKTLEAY 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ISTFLKMGILQEHPIPLSYSFIHLCFQEFFAAMSVYLEDEKGRGRGHSNCLIIDLEKTLEAY 660

Qy  661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSIESLH 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GIHGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSIESLH 720

Qy  721 CLYETRNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHV 764
          ||||||||||||||||||||||||||||||||||||||||||||
Db  721 CLYETRNKIFLTQVMAHFEEMGMCVETDMELLVCTFCIKFSRHV 764


RESULT 14
US-09-895-298-139
;  Sequence 139, Application US/09995298
;  Publication No. US20030078405A1
;  GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 47 Human Secreted Proteins
;  FILE REFERENCE: PZ035P1
;  CURRENT APPLICATION NUMBER: US/09/895,298
;  CURRENT FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: 09/591,16
;  PRIOR FILING DATE: 2000-06-09
```

```
;  PRIOR APPLICATION NUMBER: PCT/US99/29950
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: 60/113,006
;  PRIOR FILING DATE: 1998-12-18
;  PRIOR APPLICATION NUMBER: 60/112,809
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 231
;  SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 139
;  LENGTH: 442
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-895-298-139

Query Match             28.8%;  Score 2171.5;  DB 11;  Length 442;
Best Local Similarity   97.4%;  Pred. No. 8.4e-164;
Matches  414;  Conservative   4;  Mismatches   4;  Indels   3;  Gaps   1;

Qy  997 GLDTGEMSNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVEL 1056
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GLDTGEMSNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPVEL 60

Qy 1057 LCVFSPASQGDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHPPVAGSYRWPNTGLC 1116
          ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LCMPSPASQGDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHPPVAGSYRWPNTGLC 120

Qy 1117 FVMREAVTVEIEFCWWDQFLGEINPQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGH 1176
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 FVMREAVTVEIEFCWWDQFLGEINPQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGH 180

Qy 1177 VDTSLFQMAHFKEEGMLLEKPARVELHHIVLENPSFSPLGVTLLKMIHNALRFIPVTSVVL 1236
          |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 VDTSLFQVAHFKEEGMLLEKPARVELHHIVLENPSFSPLGVTLLKMIHNALRFIPVTSVVL 240

Qy 1237 LYHRVVHPEEVTFHLYLIPSDCSIRKELELCYRSPGEDQLFSEFYYGHLGSGIRLQVKDKK 1296
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LYHRVVHPEEVTFHLYLIPSDCSIRKELELCYRSPGEDQLFSEFYYGHLGSGIRLQVKDKK 300

Qy 1297 DETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVL 1356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DETLVWEALVKPGDLMPATTLIPPARISVPSPLDAPQLLHFVDQYREQLIARVTSVEVVL 360

Qy 1357 DKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKCKCKDGLYQALKETHPHLIMELM 1416
          ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db  361 DKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSWDRKCKCKDGLYQALKETHPH---SLW 417

Qy 1417 EKGSK 1421
          |:
Db  418 NSGRR 422


RESULT 15
US-10-407-866-68
;  Sequence 68, Application US/10407866
;  Publication No. US20040002593A1
;  GENERAL INFORMATION:
;  APPLICANT: Reed, John C.
;  APPLICANT: Godzik, Adam
;  APPLICANT: Reed, John C.
;  TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
;  TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
;  FILE REFERENCE: 66654-10(LJ 5755)
;  CURRENT APPLICATION NUMBER: US/10/407,866
;  CURRENT FILING DATE: 2003-04-04
;  PRIOR APPLICATION NUMBER: US 60/370,538
;  PRIOR FILING DATE: 2002-04-04
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 68
;  LENGTH: 1027
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-407-866-68
```

Query Match   18.4%; Score 1382.5; DB 12; Length 1027;
Best Local Similarity   34.4%; Pred. No. 1.9e-100;
Matches 371; Conservative 143; Mismatches 373; Indels 193; Gaps 24;

Db   898 --GDLGVLLLCEGL------RHPTCKLQTLRLGICRLGSAACBGLSVVLQANNNLRELDLS 950
            |:| ||| :| | |||:| ||| |||| | ||:| ||:|
Qy

Search completed: January 29, 2004, 13:54:37
Job time : 252.019 secs

```
Qy      8 RLACYLEFLKKEELKEFQLLIANKAHSRSSSGETP-AQPEKTSGMEVASYLVAQYGEQRA 66
          |:|| ||| ::: ::                   :  ::||:: :|:|: |::|
Db     12 RLSTYLEBLEAVELKKFKLYLGTA--TELGEGKIPWGSMEKAGPLEMAQLLITHFGPEEA 69

Qy     67 WDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPEPHLGSPSQ-PTSTAVLMPWIHELP 125
          | ||:|| ||||||                  :|      : :   ::|| |||:||||
Db     70 WRLALSTFERINRKDLWERGQREDLVRDTPPGGPSS--LGNQSTCLLEVSLVTP--RKDP 125

Qy    126 AGCTQGSERRVLRQL-PDTSGRRWREISASLLYQALPSSPDHESPSQESPNAPTSTAVLGS 185
          |::| :|| | |:|  |  || |  | :: ||  |  | ::| ::|  |: |:|:|||||
Db    126 QETYRDYAVRRKFRLMEDRNARLGECVNLSHRYTRLLLVKEHSNPMQVQ-------- 173

Qy    186 WGSPPQPSLAPREQEAPGTQWPLDETSGIYYTEIRERBREKSEKGRPPWAAVGTPPQAH 245
Db    174 ----------------------------------------------------------- 173

Qy    246 TSLQPHHHPWEPSVRESLCSTWPWKNEDFNQKFTQLILLQRPHRSQDPLVKRSWPDYVE 305
                                            ||||||  ::::|  | ::::|::|
Db    174 ------------------------------------QQLLDTGRGHARTV-------- 187

Qy    306 ENRGHL------IEIRDLFGPGLDTQE-PRIVILQGRAAGIGKSTLARQVKEAWGRGQLYGDRF 361
          ::    ::|      |::|||| || ||  :|:| |||||||||:|||||||||  :||||
Db    188 --GHQASPIKIETLFEPDEERPEPPRTVVWQGRAAGIGKSMLAHKVMLDWADGKLFQGRF 244

Qy    362 QHVFYFSCRELAQSKV-VSLAELIGKDGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
          : | | :|||||:|    |||||||:|| |||||||||:||||::|||||||||||||||
Db    245 DYLFYINCREMNQSATECSMQDLIFSCWPFPSAPLQELIRVPERLLFIIDGFDELKPSFH 304

Qy    421 EPSSELCLHWSQPQPADALLGSLLGKTIILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
          :|||:||  |:||:|  ||||| |||:||  |||:||| || |:|||||||||| ||:||
Db    305 DPQGPWCLCWEEKRPTELLLNSLIRKKLLPELSLLITTRPTALEKLHRLEHPRHVEILG 364

Qy    481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNKELWALCLVPWVSWILACTCLMQQMRKEKL 540
          | ::|::||:||:| :|   ::| | : | || ||| || :|  |:| |||||| :: |
Db    365 FSEAERKEYFYKYFHNAEQAGQVFNYVRDNEPLFTMCFVPLVCWVVCTCLQQQLEGGGLL 424

Qy    541 TLISKTTTTLCLHYLAQALQAQPLGPQL------RDLCSLAAEGIWQKKTLFSPDDLRKH 594
          | :: |    ||:|  ||:|  |:|:         |:|| | |:|:|: |  |:||:||
Db    425 RQTSRTTTAVVMLYLLSLWQPKPGAPRLQPPNQRGLCSLAADGLWNQKILFEEQDLRKH 484

Qy    595 GLDGAIISTFLKMGILQEH-PIPLSYSFIHLCFQEFFAAMSYVL--EDEKRGKHSNCIID 652
          |||| : |   |:|| :|  | || |:|  ||||:||||||||   ||:||||||  ||
Db    485 GLDGEDVSAFLNMNIFQKDINCERYVSFIHLSFQEFFAAMYYILDEGEGGAGPDQ----D 540

Qy    653 LEKTLEAYGI-HGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ- 708
          |:|:|:|  : |||    :|| | | :|:|       : | | |  |||:||| ||
Db    541 VTRLLTEYAFSERSFLALTSRFLFGLLNEETRSHLEKSLCWKVSPHIKMDLLQWIQSKAQ 600

Qy    709 ---LLLQPHSLESLHCLYETRNKTFLTQVMAHFEEWGMC--VETDMELLVCTFCIKFSRHV 764
                |:|| |||||  |:||:||||    | ||   ::  | | :|::|:|||:||||||
Db    601 SDGSTLQQGSLEFFSCLYEIQEEEFIQQALSHFQVIVVSNIASKMEHMVSSFCLKRCRSA 660

Qy    765 KKLQLI------EGRQHRSTWSPTMVVLFR--WVPVTDAYWQILFSVLKVTRNLKELD 814
          |:|:|:       | ::| |  | :| |:   || |||| |:|:|| ||| | : |
Db    661 QVLHIYGATVSADGEDRARCSAGAHTLLVQLRPERTVLLDAYSEHLAAAALCTNPNLIELS 720

Qy    815 LSGNSLSHSAVKSLCKTLRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTLTELDLSF 874
          | ||:|:  : || | :||||::||||  :| | | ::| :|||:| ||: |:|||||
Db    721 LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 780

Qy    875 NVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDLASVLSASPSLKELDLQQNNL 934
          |||  |  | || |||:  |:|:|| ::  ||  :| :|||:| |  :|||||||:|||
Db    781 NGVGFPGMMLLCEGLRHPQCRLQMIQLRKCQLESGACQEMASVLGTNPHLVELDLTGNAL 840

Qy    935 DDVGVRLLCEGLRHPACKLIRLGL------DQTTLSDEMRQELRALEQEKPQLLIF 984
          |||| |:|||||||| |||:|||:      | |:|||:|| |||:|::| |||:||
Db    841 EDLGLRLLCQGLRHPVCRLRTLWLKICRLTAAACDELASTLSVNQSIRELDLSNEL--- 897

Qy    985 SRRKPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAASHV------AQANLKLLDVS 1035
```

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 2662.52 Seconds
                                              (without alignments)
                                              11370.118 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcgggg.........atacgaaagccagcttgaa 740
Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
         1: gb_ba:*
         2: gb_htg:*
         3: gb_in:*
         4: gb_om:*
         5: gb_ov:*
         6: gb_pat:*
         7: gb_ph:*
         8: gb_pl:*
         9: gb_pr:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*
        15: em_ba:*
        16: em_fun:*
        17: em_hum:*
        18: em_in:*
        19: em_mu:*
        20: em_om:*
        21: em_or:*
        22: em_ov:*
        23: em_pat:*
        24: em_ph:*
        25: em_pl:*
        26: em_ro:*
        27: em_sts:*
        28: em_un:*
        29: em_vi:*
        30: em_htg_hum:*
        31: em_htg_inv:*
        32: em_htg_mus:*
        33: em_htg_other:*
        34: em_htg_pln:*
        35: em_htg_rod:*
        36: em_htg_mam:*
        37: em_htg_vrt:*
        38: em_sy:*
        39: em_htgo_hum:*
        40: em_htgo_mus:*
        41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	740	100.0	740	6	AR256267	AR256267 Sequence
2	740	100.0	740	6	AX082246	AX082246 Sequence
3	740	100.0	740	6	AX536208	AX536208 Sequence
4	740	100.0	740	6	AX536223	AX536223 Sequence
5	740	100.0	740	6	AF384665	AF384665 Homo sapi
6	734	99.2	770	9	AX118619	AX118619 Sequence
7	734	99.2	770	9	AF184073	AF184073 Homo sapi
8	734	99.2	782	9	AX459863	AX459863 Homo sapi
9	734	99.2	782	9	AB023416	AB023416 Homo sapi
10	733	99.1	779	9	AX017270	AX017270 Sequence
11	733	99.1	779	9	AX524974	AX524974 Sequence
12	733	99.1	779	6	BD134441	BD134441 Human nuc
13	730.8	98.8	785	6	BD057255	BD057255 Apoptois
14	713	96.4	740	9	AK000211	AK000211 Homo sapi
15	687	92.8	712	9	AF310103	AF310103 Homo sapi
16	680	91.9	772	9	BC013569	BC013569 Homo sapi
17	610	82.4	713	9	AX118641	AX118641 Sequence
18	610	82.4	713	9	AF255794	AF255794 Homo sapi
19	585	79.1	585	6	AR256268	AR256268 Sequence
20	585	79.1	585	6	AX082248	AX082248 Sequence
21	585	79.1	585	6	AX536210	AX536210 Sequence
22	448	60.5	450	12	BT007571	BT007571 Synthetic
23	437.6	59.1	768	10	BC004470	BC004470 Homo sapi
24	416.8	56.3	588	4	AB050006	AB050006 Bos tauru
25	378	51.1	405	6	AX118643	AX118643 Sequence
26	360	48.6	2821	9	AX118618	AX118618 Sequence
27	360	48.6	2821	9	AF184072	AF184072 Homo sapi
28	360	48.6	1277769	9	AC009088	AC009088 Homo sapi
29	356	48.1	356	6	AX118636	AX118636 Sequence
30	336	45.4	732	10	BC008252	BC008252 Mus muscu
31	336	45.4	770	10	AB032249	AB032249 Mus muscu
32	336	45.4	777	6	AR256273	AR256273 Sequence
33	336	45.4	777	6	AX082258	AX082258 Sequence
34	336	45.4	777	6	AX536205	AX536205 Sequence
35	336	45.4	777	6	AX536222	AX536222 Sequence
36	336	45.4	803	9	AX118637	AX118637 Sequence
37	333	45.0	701	10	AF310104	AF310104 Mus muscu
38	322	43.5	626	6	AX118621	AX118621 Sequence
39	321	43.4	340	6	AX118622	AX118622 Sequence
40	316.2	42.7	579	6	AR256274	AR256274 Sequence
41	316.2	42.7	579	6	AX082260	AX082260 Sequence
42	316	42.7	582	10	AB053165	AB053165 Rattus no
43	309.8	41.9	321	6	AX408419	AX408419 Sequence
44	293.8	39.7	579	6	AX536207	AX536207 Sequence
45	276.8	37.4	605	6	AX118639	AX118639 Sequence

ALIGNMENTS

RESULT 1
AR256267

LOCUS	AR256267	740 bp	DNA	linear	PAT	20-DEC-2002
DEFINITION	Sequence 48 from patent US 6482933.					
ACCESSION	AR256267					
VERSION	AR256267.1	GI:27305690				
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
	Unclassified.					
REFERENCE	1 (bases 1 to 740)					
AUTHORS	Bertin,J.					
TITLE	Molecules of the card-related protein family and uses thereof					
JOURNAL	Patent: US 6482933-A 48 19-NOV-2002;					
FEATURES	Location/Qualifiers					

```
RESULT 2
AX082246
LOCUS       AX082246      740 bp    DNA       linear   PAT 27-FEB-2001
DEFINITION  Sequence 48 from Patent WO0100826.
ACCESSION   AX082246
VERSION     AX082246.1  GI:13171007
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE   1
  AUTHORS   Bertin,J.
  TITLE   Novel molecules of the card-related protein family and uses thereof
  JOURNAL   Patent: WO 0100826-A 48 04-JAN-2001;
            Millennium Pharmaceuticals, Inc. (US).

```
FEATURES             Location/Qualifiers
     source          1..740
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             54..641
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAC33147.1"
                     /db_xref="GI:13171008"
                     /translation="MGRARDAILDALENLTABBLKKFKLKLLSVPLREGYGRIPRGAL
                     LSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQA
                     PPQSAAKPGLHFIDQHRAALIARVTNNVEWILDALYGKVLITDEQYQAVRABPTNPSKQA
                     KLFSFTPAWNWTCKDLLLQALRESQSYLVEDLERS"
```

```
BASE COUNT      146 a    238 c    236 g    120 t
ORIGIN
```

Query Match   100.0%; Score 740; DB 6; Length 740;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3

```
Qy      721 ATACGAAAAGCCAGCTTGAA 740
                ||||||||||||||||||||
Db      721 ATACGAAAAGCCAGCTTGAA 740
```

LOCUS   AX536208
AX536208
DEFINITION Sequence 4 from Patent WO0244354.
ACCESSION   AX536208
VERSION   AX536208.1 GI:25262599
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bertin,J.
  TITLE   Novel molecules of the card-related protein family and uses thereof
  JOURNAL   Patent: WO 0244354-A 4 06-JUN-2002;
            MILLENIUM PHARMACEUTICALS, INC. (US)

FEATURES   Location/Qualifiers
     source   1..740
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS   54..641
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57246.1"
                     /db_xref="GI:25262600"
                     /translation="MGRARDAILDALENLTABELKKFKLLLSVPLREGYGRIPRGAL
                     LSMDALDTDKLVSFYLETYGABLTANVLRDMGLQEMAGQLQAATHQQSGAAPAGIQA
                     PPQSAAKPGLRPIDQHRRAALIARVTNVEWLLDALYGKVLITDEQYQAVRAEPTNPSKMR
                     KLFSFTPAWWTCKDLILQALRESQSYLVEDLERS"

BASE COUNT   146 a   238 c   236 g   120 t
ORIGIN

Query Match   100.0%;   Score 740;   DB 6;   Length 740;
Best Local Similarity   100.0%;   Pred. No. 3.1e-107;
Matches 740;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy        1 CGGCGTCCGCGCTGCAGCGCGGGGGTGAGCGGCGGCCGGCGCGGGGGGATCCTGAGCCATGGGGGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 CGGCGTCCGCGCTGCAGCGCGGGGGTGAGCGGCGGCCGGCGCGGGGGGATCCTGAGCCATGGGGGC 60

Qy       61 GCGCGGCCGCGACGCGCCATCCTGGATGGCTGGAGAACCTGACGGCCGAGGAGCTCAAGAAGT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GCGCGGCCGCGACGCGCCATCCTGGATGGCTGGAGAACCTGACGGCCGAGGAGCTCAAGAAGT 120

Qy      121 TCAAGCTGAAGCTGCTGTCGTGGTGCCGCTGCGCGGAGGGCTTACGGCGGCCATCCCGCGCGGGGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TCAAGCTGAAGCTGCTGTCGTGGTGCCGCTGCGCGGAGGGCTTACGGCGGCCATCCCGCGCGGGGG 180

Qy      181 CGCTGCTGTCCATGGACGCGGCCTTGGACCTCACCGACACGCGGACCTCACCGACAAGCTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CGCTGCTGTCCATGGACGCGGCCTTGGACCTCACCGACACGCGGACCTCACCGACAAGCTG 240

Qy      241 CCTACGGGCCGCGAGCTCACCGCTAACGTGCTGCGGGACATGGGCCTGCCAGGAGGATGGGCCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CCTACGGGCCGCGAGCTCACCGCTAACGTGCTGCGGGACATGGGCCTGCCAGGAGGATGGGCCG 300

Qy      301 GGCAGCTGCAGGGCGCCACGCACCAGGGGGCTTCTGGAGCCGGAGGCGCCAGCTGGGATTCCAGGCCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGCAGCTGCAGGGCGCCACGCACCAGGGGGCTTCTGGAGCCGGAGGCGCCAGCTGGGATTCCAGGCCC 360

Qy      361 CTCCTCAGTGCGGCAGCCAAGCCACCACCAGGGGTCTTGGAGCCGCCGGATCCAGGCCTTA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CTCCTCAGTGCGGCAGCCAAGCCACCACCAGGGGTCTTGGAGCCGCCGGATCCAGGCCTTA 420

Qy      421 TCGCGAGGGGTCACAAACGTTGAGTGGTCGCTGCTGGATGCTCTGTACGGGAAGGTCCTGACGG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCGCGAGGGGTCACAAACGTTGAGTGGTCGCTGCTGGATGCTCTGTACGGGAAGGTCCTGACGG 480
```

RESULT 4

```
Qy      721 ATACGAAAAGCCAGCTTGAA 740
                |||||||||||||||||||
Db  ─   721 ATACGAAAAGCCAGCTTGAA 740
```

LOCUS   AX536223/c
AX536223
DEFINITION Sequence 19 from Patent WO0244354.
ACCESSION   AX536223
VERSION   AX536223.1 GI:25262610
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bertin,J.
  TITLE   Novel molecules of the card-related protein family and uses thereof
  JOURNAL   Patent: WO 0244354-A 19 06-JUN-2002;
            MILLENIUM PHARMACEUTICALS, INC. (US)

FEATURES   Location/Qualifiers
     source   1..740
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

BASE COUNT   120 a   236 c   238 g   146 t
ORIGIN

Query Match   100.0%;   Score 740;   DB 6;   Length 740;
Best Local Similarity   100.0%;   Pred. No. 3.1e-107;
Matches 740;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy        1 CGGCGTCCGGCTGCACGCGGGGGTGAGCGGCGGCCGCGGGCCGCGGGGGGATCCTGAGCCATGGGGGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      740 CGGCGTCCGGCTGCACGCGGGGGTGAGCGGCGGCCGCGGGCCGCGGGGGGATCCTGAGCCATGGGGGC 681

Qy       61 GCGCGGCCGCGACGCGCCATCCTGGATGGCTGGAGAACCTGACGGCCGAGGAGCTCAAGAAGT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      680 GCGCGGCCGCGACGCGCCATCCTGGATGGCTGGAGAACCTGACGGCCGAGGAGCTCAAGAAGT 621

Qy      121 TCAAGCTGAAGCTGCTGTCGTGGTGCCGCTGCGCGGAGGGCTACGGCGGCCATCCCGCGCGGGGG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      620 TCAAGCTGAAGCTGCTGTCGTGGTGCCGCTGCGCGGAGGGCTACGGCGGCCATCCCGCGCGGGGG 561

Qy      181 CGCTGCTGTCCATGGACGCGGCCTTGGACCTCACCGACACGCGGACCTCACCGACAAGCTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      560 CGCTGCTGTCCATGGACGCGGCCTTGGACCTCACCGACACGCGGACCTCACCGACAAGCTG 501

Qy      241 CCTACGGGCCGCGAGCTCACCGCTAACGTGCTGCGGGACATGGGCCTGCCAGGAGGATGGGCCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      500 CCTACGGGCCGCGAGCTCACCGCTAACGTGCTGCGGGACATGGGCCTGCCAGGAGGATGGGCCG 441

Qy      301 GGCAGCTGCAGGGCGCCACGCACCAGGGGCTTCTGGAGCCGCGCCAGCTGGGATCCAGGCCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGCAGCTGCAGGGCGCCACGCACCAGGGGCTTCTGGAGCCGCGCCAGCTGGGATCCAGGCCC 360
```

RESULT 5

LOCUS   AF384665   740 bp   mRNA   linear   PRI 19-JUN-2001
DEFINITION   Homo sapiens caspase recruitment domain protein 5 mRNA, complete
                cds.
ACCESSION   AF384665
VERSION   AF384665.1 GI:14488058
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 740)
  AUTHORS   Bertin,J.
  TITLE   CARD5 Protein is a CARD/PYRIN family member that is involved in
                apoptosis signal transduction
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 740)
  AUTHORS   Bertin,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals
                Inc., 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES   Location/Qualifiers
     source   1..740
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS   54..641
                     /note="CARD5; CARD/PYRIN family member; bipartite protein
                     comprised of an N-terminal PYRIN domain and a C-terminal
                     CARD domain; involved in apoptosis signal transduction"
                     /codon_start=1
                     /product="caspase recruitment domain protein 5"
                     /protein_id="AAK63850.1"
                     /db_xref="GI:14488059"
                     /translation="MGRARDAILDALENLTPAEELKKFKLKLLSVPLREGYGRIPRGAL
                     LSMDALDLTDKLVSFYVLETYGAELTANVLRDMGIQEMAGQLQAATHQGSGAAPAGIQA
                     PPQSAAKPGLHFIDQHRAALIARVTNVEWLLDDALYGKVLTDEQYQAVRAEPTNPSKMR
                     KLFSFTPAWNWTCKDLLQALRESQSYLVEDLERS"

BASE COUNT   146 a   238 c   236 g   120 t
ORIGIN

Query Match   100.0%; Score 740; DB 9; Length 740;
Best Local Similarity   100.0%;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CGGCGTCCGGCTGCAGGGGGTGAGGCGGGGTGAGCCGGCGCAGGCGGCGGGGATCCTGGACCGATGGGGCGC 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGGCGTCCGGCTGCAGGGGGTGAGGCGGGGTGAGCCGGCGCAGGCGGCGGGGATCCTGGACCGATGGGGCGC 60

[sequence alignment continues]

RESULT 6

LOCUS   AX118619   770 bp   DNA   linear   PAT 11-MAY-2001
DEFINITION   Sequence 2 from Patent WO0129235.
ACCESSION   AX118619
VERSION   AX118619.1 GI:14035570
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Vertino,P.M.
  TITLE   Tmsl compositions and methods of use
  JOURNAL   Patent: WO 0129235-A 2 26-APR-2001;
                Emory University (US)

Left column

```
FEATURES             Location/Qualifiers
     source          1..770
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             75..662
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAC38471.1"
                     /db_xref="GI:14035571"
                     /translation="MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGAL
                     LSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAPAGIQA
                     PQGSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR
                     KLFSFTPAWWTCKDLLLQALRESQSYLVEDLERS"
BASE COUNT     157 a     247 c     241 g     125 t
ORIGIN

Query Match      99.2%;  Score 734;  DB 6;  Length 770;
Best Local Similarity   100.0%;   Pred. No. 2.8e-106;
Matches 734;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     7 CGGGCTGCAGCGGGGGTGAGCGGCCAGCGGCGCGGGGGGATCCTTGGACGCCATGGGGCGCCGGCGC 66
         ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28 CGGGCTGCACCGGGGGTGAGCGGCCAGCGGCGCGGGGGGATCCTTGGACGCCATGGGGCGCCGGCGC 87

Qy    67 GCGAGGCCATCCTGGATGCGGCTGGAGAAACCTGACCCGCGGGGGGCTCAAGAAGTTCAAGC 126
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88 GCGAGGCCATCCTGGATGCGGCTGGAGAAACCTGACCCGCGGGGGGCTCAAGAAGTTCAAGC 147

Qy   127 TGAAGCTGCTGTCGGTGCCGCTGCGCGAGGGCTACGGCCGCATCCCGCGCGGGCTGC 186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148 TGAAGCTGCTGTCGGTGCCGCTGCGCGAGGGCTACGGCCGCATCCCGCGCGGGCTGC 207

Qy   187 TGTCCATGGACGCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 246
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 TGTCCATGGACGCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 267

Qy   247 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCCTGCAGGAGATGGCCGGGCAGC 306
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCCTGCAGGAGATGGCCGGGCAGC 327

Qy   307 TGCAGGGCGGCCACGCACCAGGGCTCTTGGAGCGCCAGCTGGGATCCAGGCCCCTCCTC 366
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 TGCAGGGCGGCCACGCACCAGGGCTCTTGGAGCGCCAGCTGGGATCCAGGCCCCTCCTC 387

Qy   367 AGTCCGGCAGCCAGCCAGGCCTGCACTTTATAGACCAGCACCGGGGCTTATCGGCGA 426
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 AGTCCGGCAGCCAGCCAGGCCTGCACTTTATAGACCAGCACCGGGGCTTATCGGCGA 447

Qy   427 GGGGTCACAAAACGTTGAGTGGCTGCTGGATGTTCTGTTGTACGGGAAGGTCCTGACGGATGAGC 486
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 GGGGTCACAAAACGTTGAGTGGCTGCTGGATGTTCTGTTGTACGGGAAGGTCCTGACGGATGAGC 507

Qy   487 AGTACCAGGCAGTGCGGGCGGCCGAGCCCAACCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 546
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 AGTACCAGGCAGTGCGGGCGGCCGAGCCCAACCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 567

Qy   547 TCACACCAGCCTGGAACTGGACCTGCGAAGGACTTGCTCTTCCTTCCAGGCGCCCTAAGGGAGTCCC 606
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 TCACACCAGCCTGGAACTGGACCTGCGAAGGACTTGCTCTTCCTTCCAGGCGCCCTAAGGGAGTCCC 627

Qy   607 AGTCCTTACCTGGTGAGGAGGGACCTGGACGTGAGGCTCCTCCAGCGACCAACACTCCGGT 666
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 AGTCCTTACCTGGTGAGGAGGGACCTGGACGTGAGGCTCCTCCAGCGACCAACACTCCGGT 687

Qy   667 CAGCCCCTGGCAATCCCACCAAAATCATCCTTGAATCTTGATCTTTTTTATACACAATATACGA 726
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 CAGCCCCTGGCAATCCCACCAAAATCATCCTTGAATCTTGATCTTTTTTATACACAATATACGA 747

Qy   727 AAAGGCCAGCTTGAA 740
         ||||||||||||||
Db   748 AAAGGCCAGCTTGAA 761
```

Right column

```
RESULT 7
AF184073
LOCUS       AF184073     770 bp    mRNA     linear   PRI 02-FEB-2001
DEFINITION  Homo sapiens target of methylation-induced silencing 1 (TMS1) mRNA,
            complete cds.
ACCESSION   AF184073
VERSION     AF184073.1  GI:9863863
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.
            and Vertino,P.M.
  TITLE     TMS1, a novel proapoptotic caspase recruitment domain protein, is a
            target of methylation-induced gene silencing in human breast
            cancers
  JOURNAL   Cancer Res. 60 (22), 6236-6242 (2000)
   MEDLINE  20552139
   PUBMED   11103776
REFERENCE   2  (bases 1 to 770)
  AUTHORS   McConnell,B.B. and Vertino,P.M.
  TITLE     Activation of a caspase-9-mediated apoptotic pathway by subcellular
            redistribution of the novel caspase recruitment domain protein TMS1
  JOURNAL   Cancer Res. 60 (22), 6243-6247 (2000)
   MEDLINE  20552140
   PUBMED   11103777
REFERENCE   3  (bases 1 to 770)
  AUTHORS   Vertino,P.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-SEP-1999) Radiation Oncology, Emory University School
            of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
FEATURES             Location/Qualifiers
     source          1..770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1..770
                     /gene="TMS1"
                     /note="involved in apoptosis"
                     /function="contains caspase-recruitment domain; CARD protein;
                     target of methylation-mediated gene silencing in Homo
                     sapien tumors"
     CDS             75..662
                     /gene="TMS1"
                     /codon_start=1
                     /product="target of methylation-induced silencing 1"
                     /protein_id="AAG01188.1"
                     /db_xref="GI:9863864"
                     /translation="MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGAL
                     LSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAPAGIQA
                     PQGSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR
                     KLFSFTPAWWTCKDLLLQALRESQSYLVEDLERS"
BASE COUNT     157 a     247 c     241 g     125 t
ORIGIN

Query Match      99.2%;  Score 734;  DB 9;  Length 770;
Best Local Similarity   100.0%;   Pred. No. 2.8e-106;
Matches 734;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     7 CGGGCTGCAGCGGGGGTGAGCGGCCAGCGGCGCGGGGGGATCCTTGGACGCCATGGGGCGCCGGCGC 66
         ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28 CGGGCTGCACCGGGGGTGAGCGGCCAGCGGCGCGGGGGGATCCTTGGACGCCATGGGGCGCCGGCGC 87

Qy    67 GCGAGGCCATCCTGGATGCGGCTGGAGAAACCTGACCCGCGGGGGGCTCAAGAAGTTCAAGC 126
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88 GCGAGGCCATCCTGGATGCGGCTGGAGAAACCTGACCCGCGGGGGGCTCAAGAAGTTCAAGC 147
```

```
RESULT 8
AX459863

LOCUS       AX459863      782 bp    DNA      linear    PAT 08-JUL-2002
DEFINITION  Sequence 6 from Patent WO0240668.
ACCESSION   AX459863
VERSION     AX459863.1  GI:21725640
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tschopp,J. and Martinon,F.
  TITLE     Proteins and dna sequences underlying these proteins used for
            treating inflammations
  JOURNAL   Patent: WO 0240668-A 6 23-MAY-2002;
            Apotech Research and Development Ltd. (CH)
FEATURES             Location/Qualifiers
     source          1..782
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Pycard.cdna"

BASE COUNT      158 a    251 c    246 g    127 t
ORIGIN
Query Match       99.2%;  Score 734;  DB 6;  Length 782;
Best Local Similarity  100.0%;  Pred. No. 2.8e-106;

Matches   734;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy    127 TGAAGCTGCTGCTGTCGGTGCGCGTGCGGCGGGCTACGCGGCGCTGC  186
          |||||||||||||||||||||||||||||||||||||||||||||||
Db      7 CGGCTGCAGCGGGGGTGACGCGCGGGCGGGATCCTGGAGCCATGGGGCGGGCGC  66

Qy    187 TGTCCATGGACGGCCTTGAGCTCACCGACAAGCTGGTTCAGCTTCACGTTC...  246
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 GCGACGGCCATCCTGGAGATGCGGCTGGAGAGAGGGCTACGCGGCGCTGAGCTCAAGG  126
```

(sequence alignment blocks continue for Qy/Db pairs from positions 127 through 761)

```
Qy    748 AAAGCCAGCTTGAA  761
          ||||||||||||||
Db    748 AAAGCCAGCTTGAA  747
```

```
RESULT 9
AB023416

LOCUS       AB023416      782 bp    mRNA     linear    PRI 18-OCT-2000
DEFINITION  Homo sapiens ASC mRNA for apoptosis-associated speck-like protein
            containing a CARD, complete cds.
ACCESSION   AB023416
VERSION     AB023416.2  GI:10801601
KEYWORDS    ASC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
  AUTHORS   Masumoto,J., Taniguchi,S., Ayukawa,K., Sarvotham,H., Kishino,T.,
            Niikawa,N., Hidaka,E., Katsuyama,T., Higuchi,T. and Sagara,J.
  TITLE     ASC, a novel 22-kDa protein, aggregates during apoptosis of human
            promyelocytic leukemia HL-60 cells
  JOURNAL   J. Biol. Chem. 274 (48), 33835-33838 (1999)
```

```
Qy    727 AAAGCCAGCTTGAA  740
          ||||||||||||||
Db    760 AAAGCCAGCTTGAA  773
```

MEDLINE   20036508
PUBMED   10567338
REFERENCE 2 (bases 1 to 782)
AUTHORS   Masumoto,J., Sagara,J. and Taniguchi,S.
TITLE   Direct Submission
JOURNAL   Submitted (04-FEB-1999) Junya Masumoto, Shinshu University School
           of Medicine, Research Center on Aging and Adaptation; Asahi 3-1-1,
           Matsumoto, Nagano 390-8621, Japan
           (E-mail:masumoto@sch.md.shinshu-u.ac.jp, Tel:81-263-37-2723,
           Fax:81-263-37-2724)
COMMENT   On Oct 14, 2000 this sequence version replaced gi:6482371.
           Sequence updated (29-Nov-1999).
FEATURES   Location/Qualifiers
     source   1..782
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /map="16p11.2-12"
     gene   1..782
                     /gene="ASC"
     CDS   87..674
                     /gene="ASC"
                     /codon_start=1
                     /product="apoptosis-associated speck-like protein
                     containing a CARD"
                     /protein_id="BAA87339.2"
                     /db_xref="GI:10801602"
                     /translation="MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGAL
                     LSMDALDLTDKLVSFYLETYGAELTANYLRDMGLQEWAGQLQAATHQGSGAAPAGIQA
                     PQSAAKPGLHFIDQHRRALIRAVTNVEWLLDALYGKVLTDEQYQAVRABPTNPSKMR
                     KLFSFPAWNWTCKDLLLQALRESQSYLVEDLERS"
BASE COUNT   158 a   251 c   246 g   127 t
ORIGIN

Query Match   99.2%;   Score 734; DB 9; Length 782;
Best Local Similarity   100.0%;   Pred. No. 2.8e-106;
Matches 734; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   7 CGGGCTGCAGCGGGGGTGAGCGGGCGGCGGCGGGGGGGATCCTTGGAGCCATGGGGCGCGC 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   40 CGGGCTGCACCGGGGGTGAGCGGGCGGCGCGCCGGGGATCCTTGGAGCCATGGGGCGCGC 99

Qy   67 GCGACGCCCATCCTGGATGCGCTGGAGAACCTGACCGACCCGGAGGAGCTCAAGAAGTTCAAGC 126
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 GCGACGCCCATCCTGGATGCGCTGGAGAACCTGACCGACCCGGAGGAGCTCAAGAAGTTCAAGC 159

Qy   127 TGAAGCTGTCGTTCGGTCCGCTCCCGCGCGAGGGGCTACGGGCGCCATCCCGCGCGGGCTGC 186
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 TGAAGCTGCTGTTCGGTGCGCTGCCGCGCGAGGGGCCATCCGCGCGGGGCTACGGGCGCCTGC 219

Qy   187 TGTCCATGGACGCCCTTGGACCTCACCGACACGAAGCTGGTCAGCTTCTACCTGGAGACCTACG 246
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 TGTCCATGAACGCCCTTGGACCTCACCGACACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 279

Qy   247 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGGCCTCAGGAGATGGCCGGGCAGC 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGGCTCAGGAGAGATGGCCGGGCAGC 339

Qy   307 TGCAGGCGGCCACCCACCCAGGGGCCTCTGGGACCCCAGGCTGGGGATCCAGGCCCCTCCTC 366
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340 TGCAGGCGGCCACCCAGCCAGGGTCTGGGGATCCAGGCCCCTCCTC 399

Qy   367 AGTCGGGACCAAGCCAGGCCTGCACTTTATAGACCAGCACACGCGGCCCTTATGCGGCGA 426
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   400 AGTCGGGACCAAGCCAGGCCTGCACTTTATAGACCAGCACACGCGGCCCTTATCGCGGT 459

Qy   427 GGGGTCACAAACGTTGAGTGGCTGCTGGAGAAGGTTCCTGACGGATGAGC 486
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   460 GGGGTCACAAACGTTGAGTGGCTGCTTGGAATGGCTTCTCTGACGGAATGAGC 519

Qy   487 AGTACCAGGCGAGTGCGGGGCCGGAGCCCACCAACCCAAGCAAGAGATGCGGAAGGCTCTTCAGTT 546
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   520 AGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGAGATGCGGAAAGCTCTTCAGTT 579

Qy   547 TCACACCAGGCCTGGAACTGGACCTGCAAGGACACTTGCTCCTCCAGGCGCCTAAGGGAGTCC 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   580 TCACACCAGGCCTGGAACTGGACCTGCAAGGACACTTGCTCCTCCAGGCGCCTAAGGGAGTCC-639

Qy   607 AGTCCTTACCTGGGTGGAGGAGCCAGGCGAGGAGGCTCCTTCCCTCCAGCGACACACTCCGGT 666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   640 AGTCTTACCTGGGTGGAGGAGGAGCCAGGCGAGGGAGGCTCCTTCCCTCCAGCGACACACTCCGGT 699

Qy   667 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAAATCTGATCTTTTATTACACAATATATACGA 726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   700 CAGCCCCTGGCAAATCCCACCAAAATCATCCTGAAATCTGATCTTTTTATTACACAATATATACGA 759

Qy   727 AAAGCCAGCTTGAA 740
            |||||||||||||
Db   760 AAAGCCAGCTTGAA 773

RESULT 10
LOCUS   AX017270   779 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION Sequence 21 from Patent WO9947669.
ACCESSION   AX017270
VERSION   AX017270.1 GI:10042188
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
            Pilarsky,C.
  TITLE   Human nucleic acid sequences from tissue of breast tumors
  JOURNAL   Patent: WO 9947669-A 21 23-SEP-1999;
            SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
            BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
            (DE); PILARSKY CHRISTIAN (DE)
FEATURES   Location/Qualifiers
     source   1..779
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

BASE COUNT   149 a   252 c   251 g   127 t

Query Match   99.1%;   Score 733; DB 6; Length 779;
Best Local Similarity   100.0%;   Pred. No. 4e-106;
Matches 733; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   7 CGGGCTGCAGCGGGGGTGAGCGGGCGGCGCGCCGGGGATCCTTGGAGCCATGGGGCGCGC 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   47 CGGGCTGCACCGGGGGTGAGCGGGCGGCGCGCCGGGGATCCTTGGAGCCATGGGGCGCGC 106

Qy   67 GCGACGCCCATCCTGGATGCGCTGGAGAACCTGACCCCGGAGGAGCTCAAGAAGTTCAAGC 126
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   107 GCGACGCCCATCCTGGATGCGCTGGAGAACCTGACCCCGGAGGAGCTCAAGAAGTTCAAGC 166

Qy   127 TGAAGCTGCTGTTCGGTGCGCTGCCGCGCGAGGGGCCATCCCGCGCGGGGCTGC 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   167 TGAAGCTGCTGTTCGGTGCGCTGCCGCGCGAGGGGCCATCCCGCGCGGGGCTGC 226

Qy   187 TGTCCATGAACGCCCTTGGACCTCACCGACACAAGCTGTCAGCTTCTACCTGGAGACCTACG 246
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   227 TGTCCATGAACGCCCTTGGACCTCACCGACACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 286

Qy   247 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCCTCAGGAGAGATGGCCGGGCAGC 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   287 GCGCCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCCTCAGGAGAGATGGCCGGGCAGC 346

Qy   307 TGCAGGCGGCCACCCAGCCAGGGCTCTGGAGCCCCAGCTCGGGGGATCCAGGCCCCTCCTC 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   347 TGCAGGCGGCCACGCACCAGGGCTCTGGAGCCCCAGCTCGGGGGATCCAGGCCCCTCCTC 406

RESULT 11

```
LOCUS       AX524974               779 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 21 from Patent EP1236799.
ACCESSION   AX524974
VERSION     AX524974.1  GI:25170056
KEYWORDS
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and
            Rosenthal,A.
  TITLE     Human nucleic acid sequences derived from breast tumor tissue
  JOURNAL   Patent: EP 1236799-A 21 04-SEP-2002;
            metaGen Pharmaceuticals GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..779
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      149 a    252 c    251 g    127 t
ORIGIN
```

Query Match   99.1%; Score 733; DB 6; Length 779;
Best Local Similarity   100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      7 CGGCTGCAGCGGGGTGAGCGGCGCGGGGATCCTGAGCGGCCATGAGGGGCGGCGGGC 66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47 CGGCTGCAGCGGGGTGAGCGGCGCGGGGATCCTGAGCGGCCATGAGGGGCGGCGGGC 106

Qy     67 GCGACGCCATCCTGGATGGCGTGAGAACCTGACGCGCCGAGGAGCTCAAGAAGTTCAAGC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107 GCGACGCCATCCTGGATGGCGTGAGAACCTGACGCGCCGAGGAGCTCAAGAAGTTCAAGC 166

Qy    127 TGAAGCTGCTGTCGGTGCCGCTGCCGGCGAGGGCTACGGCGGCGCATCCCGCGGGCGGCTTGC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    167 TGAAGCTGCTGTCGGTGCCGCTGCCGGCGAGGGCTACGGCGGCGCATCCCGCGGGCGGCTTGC 226

Qy    187 TGTCCATGGACGCCTTGGACCTCACCGACAAAGTGGTCAGCTTCTTACCTGGAGACCTACG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 TGTCCATGGACGCCTTGGACCTCACCGACAAAGTGGTCAGCTTCTTACCTGGAGACCTACG 286

Qy    247 GCGCCGGAGCTCACCGCCTAACGTTGTGTGCGGGGACATGGGCCGGGCAGC 306
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287 GCGCCGGAGCTCACCGCCTAACGTTGTGTGCGGGGACATGGGCCGGGCAGC 346

Qy    307 TGCAGGCGCGGGCCACCGGACCAGGGCTCTGGACGCCGGCCCAGGTGGATCCAGGCCGTC 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 TGCAGGCGCGGGCCACCGGACCAGGGCTCTGGACGCCGGCCCAGGTGGATCCAGGCCGTC 406

Qy    367 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACAGACACACGGGCGCGCTTATCGGGCGGCGA 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACAGACACACGGGCGCGCTTATCGGGCGGCGA 466

Qy    427 GGGTCACAAAACGTTGAGTGGTGGCTGCTGCTGGACGGGAGGTCCTGGACGGATGGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 GGGTCACAAAACGTTGAGTGGTGGCTGCTGCTGGACGGGAGGTCCTGGACGGATGGAGC 526

Qy    487 AGTACCAGGCCAGTGCGGGCCGACCCAGCCCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    527 AGTACCAGGCCAGTGCGGGCCGACCCAGCCCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 586

Qy    547 TCACACCAGCCTGGAACTGGACTTGCTCCTCCTCCAGGACTTGCTCCTCCCAGGACAACACTCCGGT 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    587 TCACACCAGCCTGGAACTGGACTTGCTCCTCCTCCAGGACTTGCTCCTCCCAGGACAACACTCCGGT 646

Qy    607 AGTCCTACCTGGTGGAGGGACCTGAGGGGCTCCTTCCTCCCAGCAACACTCCGGT 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    647 AGTCCTACCTGGTGGAGGGACCTGAGGGGCTCCTTCCTCCCAGCAACACTCCGGT 706

Qy    667 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTATACACAATATATACGA 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    707 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTATACACAATATATACGA 766

Qy    727 AAAGCCAGCTTGA 739
          |||||||||||||
Db    767 AAAGCCAGCTTGA 779
```

RESULT 12

```
LOCUS       BD134441               779 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION   BD134441
VERSION     BD134441.1  GI:23229386
KEYWORDS    JP 2002506643-A/19.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Speft,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
            Rosenthal,A.
  TITLE     Human nucleic acid sequence originating in mammary tumor tissue
  JOURNAL   Patent: JP 2002506643-A 19 05-MAR-2002;
            METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002506643-A/19
            PD   05-MAR-2002
            PF   19-MAR-1999 JP 2000536852
            PF   20-MAR-1998 DB 198 13 839.3
            PI   THOMAS SPEFT,BERND HINTZMAN,SHCMITI ARMIN,CHRISTIAN PIRARSKI,
            PI   DUHL EDGAR,
            PI   ANDRE ROSENTHAL
            PC   C12N15/09,A61K31/711,A61K38/00,A61K48/00,C07K14/47,
            PC   C07K16/18,
            PC   C12N1/19,C12N15/10,C12N15/00,A61K37/02,C12N5/00 CC
            nucleic acid sequence originating in mammary tumor CC
            tissue
            FH   Key             Location/Qualifiers
            FT   source          1..779
                                 /organism='Homo sapiens (human)',
FEATURES             Location/Qualifiers
```

```
Db    227 TGTCCATGGACGCCTTGGACCTCACCGACAAAGTGGTCAGCTTCTTACCTGGAGACCTACG 286
Qy    247 GCGCCGGAGCTCACCGCCTAACGTTGTGTGCGGGGACATGGGCCGGGCAGC 306
Db    287 GCGCCGGAGCTCACCGCCTAACGTTGTGTGCGGGGACATGGGCCGGGCAGC 346
Qy    307 TGCAGGCGCGGGCCACCGGACCAGGGCTCTGGACGCCGGCCCAGGTGGATCCAGGCCGTC 366
Db    347 TGCAGGCGCGGGCCACCGGACCAGGGCTCTGGACGCCGGCCCAGGTGGATCCAGGCCGTC 406
Qy    367 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACAGACACACGGGCGCGCTTATCGGGCGGCGA 426
Db    407 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACAGACACACGGGCGCGCTTATCGGGCGGCGA 466
Qy    427 GGGTCACAAAACGTTGAGTGGTGGCTGCTGCTGGACGGGAGGTCCTGGACGGATGGAGC 486
Db    467 GGGTCACAAAACGTTGAGTGGTGGCTGCTGCTGGACGGGAGGTCCTGGACGGATGGAGC 526
Qy    487 AGTACCAGGCCAGTGCGGGCCGACCCAGCCCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 546
Db    527 AGTACCAGGCCAGTGCGGGCCGACCCAGCCCAAACCCAAGCAAGATGGGGAAGCTCTTCAGTT 586
Qy    547 TCACACCAGCCTGGAACTGGACTTGCTCCTCCTCCAGGACTTGCTCCTCCCAGGACAACACTCCGGT 606
Db    587 TCACACCAGCCTGGAACTGGACTTGCTCCTCCTCCAGGACTTGCTCCTCCCAGGACAACACTCCGGT 646
Qy    607 AGTCCTACCTGGTGGAGGGACCTGAGGGGCTCCTTCCTCCCAGCAACACTCCGGT 666
Db    647 AGTCCTACCTGGTGGAGGGACCTGAGGGGCTCCTTCCTCCCAGCAACACTCCGGT 706
Qy    667 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTATACACAATATATACGA 726
Db    707 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTATACACAATATATACGA 766
Qy    727 AAAGCCAGCTTGA 739
Db    767 AAAGCCAGCTTGA 779
```

RESULT 13

LOCUS   BD057255   785 bp   DNA   linear   PAT 27-AUG-2002
DEFINITION Apoptosis-related protein, its antibody and its DNA.
ACCESSION   BD057255
VERSION   BD057255.1 GI:22602861
KEYWORDS   JP 2001275681-A/1.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 785)
  AUTHORS   Masumoto,J., Sagara,J. and Taniguchi,S.
  TITLE   Apoptosis-related protein, its antibody and its DNA
  JOURNAL   Patent: JP 2001275681-A 1 09-OCT-2001;
            JUNYA MASUMOTO,JUNJI AIRA,SHUNICHIRO TANIGUCHI, MEDICAL &
            BIOLOGICAL LABORATORIES CO LTD
  COMMENT   OS   Homo sapiens (human)
            PN   JP 2001275681-A/1
            PD   31-MAR-2000 JP 200098204
            PF   09-OCT-2001
            PI   JUNYA MASUMOTO,JUNJI SAGARA, SHUNICHIRO TANIGUCHI PC
            CI2N15/09,C07K14/82,C07K16/32//C12P21/02,C12P21/08,C12N15/00 CC
            FH   Key   Location/Qualifiers.
            FH
FEATURES   Location/Qualifiers
     source   1..785
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT   160 a   252 c   247 g   126 t
ORIGIN

Query Match   98.8%; Score 730.8; DB 6; Length 785;
Best Local Similarity   99.7%; Pred. No. 8.8e-106;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   7 CGGCTGCAGCGCGGGGTGAGCGCGGCGCAGGCGCAGGCGCAGGCGCAGGCCCATGGGGCGC 66
Db   40 CGGCTGCAGCGCGGGGTGAGCGCGGCGCAGGCGCAGGCGCAGGCGCAGGCCCATGGGGCGC 99

Qy   127 TGAAGCTGCTGTCGGTGCCCGCTAAACGTGCTCACCGAC... 186
Db   100 ...

source   1..779
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
BASE COUNT   149 a   252 c   251 g   127 t
ORIGIN

Query Match   99.1%; Score 733; DB 6; Length 779;
Best Local Similarity   100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   7 CGGGCTGCAGCGCGGGGTGAGCGGGCGGCGGGGGATCCTGGACCATGGGGCGCGC 66
Db   47 CGGGCTGCAGCGCGGGGTGAGCGGGCGGCGGGGGATCCTGGACCATGGGGCGCGC 106

Qy   67 GCGACGGCCACCCTGGATGCGCGTGGAGAGAACCTGACCCGCGGGGGCAGC 126
Db   107 ...

Qy   727 AAAGGCAGCTTGA 739
Db   767 AAAGGCAGCTTGA 779

Db   700 CAGCCCCTTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTATACACAATAATACGA 759

Qy   727 AAAGCCAGCTTGAA 740
                |||||||||||||
Db   760 AAAGCCAGCTTGAA 773

RESULT 14

LOCUS   AK000211   740 bp   mRNA   linear   PRI 22-FEB-2000
DEFINITION   Homo sapiens cDNA FLJ20204 fis, clone COLF1505.
ACCESSION   AK000211
VERSION   AK000211.1 GI:7020146
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
                Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
                Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE   NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 740)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
                Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
                University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1,
                Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna1@ims.u-tokyo.ac.jp,
                Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
                International Trade and Industry of Japan; cDNA full insert
                sequencing: Research Association for Biotechnology; cDNA library
                construction, 5'- & 3'-end one pass sequencing: Depardent of
                Virology and Human Genome Center, Institute of Medical Science,
                University of Tokyo (partly supported by Science and Technology
                Agency).
FEATURES   Location/Qualifiers
     source   1..740
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COLF1505"
                     /tissue_type="colon mucosa"
                     /clone_lib="ColF"
                     /note="cloning vector pME18SFL3"
     CDS   177..626
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA91012.1"
                     /db_xref="GI:7020147"
                     /translation="MDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQAATH
                     QGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALIARVNVEWLLDALYGKVLTDEQYQA
                     VRAEPINPSKMRKLFSFTPAWNTCKDLLLQALRBSQSYLVEDLERS"
BASE COUNT   161 a   231 c   230 g   118 t
ORIGIN

Query Match   96.4%; Score 713; DB 9; Length 740;
Best Local Similarity 99.7%; Pred. No. 5.7e-103;
Matches   725; Conservative   0; Mismatches   0; Indels   2; Gaps   1;

Qy   14 AGCGGGGGTGAGCGCGGCGGGCGGGCGGGCCGGGGGGGGGGGGCCGGGGGGGGGGGGGCCGAGCGC 73
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1 AGCGGGGGTGAGCGCGGCGGCCGCGGGGATCCTTGGAGCCGCCATGGGGCG-GGGGGGGCCGAGCGC 58

Qy   74 CATCCTGGATCGCTGGAGAACTTGACCCCGAGGGAGCTCAAGAAGTTCAAGCTGAAGCT 133
            ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   59 CATCCTGGATCGCTGGAGAAACCTGACCCCGAGGGAGCTCAAGAAGAGTTCAAGCTGAAGCT 118

Qy   134 GCTGTCGGTGTCCGCTGCCGAGGGGCTACGGGCCGGGGCAATCCCGCGGGGGGGGGCCGCTGCGTGTCCAT 193
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 ...

RESULT 15

LOCUS   AF310103   712 bp   mRNA   linear   PRI 04-MAY-2001
DEFINITION   Homo sapiens PYCARD mRNA, complete cds.
ACCESSION   AF310103
VERSION   AF310103.1 GI:11096298
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 712)
  AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
  TITLE   The pyrin domain: a possible member of the death domain-fold family
                implicated in apoptosis and inflammation
  JOURNAL   Curr. Biol. 11 (4), R118-R120 (2001)
   MEDLINE   21148093
    PUBMED   11250163
REFERENCE   2 (bases 1 to 712)
  AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (28-SEP-2000) Institute of Biochemistry, University of
                Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
FEATURES   Location/Qualifiers
     source   1..712
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS   1..588
                     /note="contains PYD and CARD domains"

Db   119 GCTGTGTCGGTGCCGGCTGCGGCGGGGCGGGAGGGGCTACGGGGCGGGGGGGGCCCATCCCGCGGGGGGGGGGGGGCCGCTGCGTGTCCAT 178

Qy   194 GGACGGCCTTGGGACCTCACCGACAAGGCCAAGGCTGGTCAGCTTCAGCTTCTACCTGGAGACCTACGGCGCCGA 253
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 GCTCACCCGCTAAGGTGCGTCGCGCGACACATGGGCCGTCTGCAGAGAGATGGCCGGCCAGCTTGCCAGGC 298

Qy   254 AGCTCACCCGCTAAGGTGCGTCGCGCGGACAATGGGCCTTGGCAGAGAGATGGCCGGCCAGCTTGCCAGGC 313
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 GCTCACCCGCTAAGGTGCGTCGCGCGGACAATGGGCCTTGGCAGAGAGATGGCCGGCCAGCTTGCCAGGC 313

Qy   314 GGCCAGCGCACCAGGGCTCTTGGAGCCGCGCGGGTCGGGATCAGGCCCCTCCTCAGTCGGC 373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 GGCCAGCGCACCAGGGCTCTTGGAGCCGCGCGGGTCGGGATCCAGGCCCCTCCTCAGTCGGC 358

Qy   374 AGCCAAGCCCAGGGCCTGCACTTTATTGAGCAGCACCGGGCCGCCCTTATCGCCGGGTCAC 433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 AGCCAAGCCCAGGGCCTGCACTTTATTGAGCAGCACCGGGCCGCCCTTATCGCCGGGTCAC 418

Qy   434 AAACGTTGAGTGGCTGCTGGGATGCTTCGTCTACCGGGAAGGTCCTGACGGATGAGCAGTA 493
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 AAACGTTGAGTGGCTGCTGGGATGCTTCGTCTACCGGGAAGGTCCTGACGGATGAGCAGTACCA 478

Qy   494 GGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGATGCGGAAGCTCTTCAGTTTCACAC 553
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 GGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACC 538

Qy   554 AGCCTGGAACTGGACCTGCAAGGACTTCCTCCCCCAGGGCCCTAAGGGAGTCCCAGTTCCTA 613
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   539 AGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGGCCCTAAGGGAGTCCCAGTTCCTA 598

Qy   614 CCTGGTTGAGGAGCCTGGGAGCTCCTTCCCAGCCAACAACACTCCGGTGCAGCGCCC 673
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 CCTGGTTGAGGAGCCTGGGAGCTCCTTCCCAGCCAACAACACTCCGGTGCAGCCCC 658

Qy   674 TGGGAATCCCACCCAAATCATCCTGAAATCTGATCTTTTTATACACAATATACGAAAAGCCA 733
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   659 TGGGAATCCCACCCAAATCATCCTGAAATCTGATCTTTTTATACACAATATACGAAAAGCCA 718

Qy   734 GCTTGAA 740
            |||||||
Db   719 GCTTGAA 725

```
                    /codon_start=1
                    /product="PYCARD"
                    /protein_id="AAG30286.1"
                    /db_xref="GI:11096299"
                    /translation="MGRARDAILDALENLTAEELKKFKLKLLSVPLRBGYGRIPRGAL
                    LSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGDLQAAATHQGSGAPAGIQA
                    PQSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTIDEQYQAVRAEPTNPSKMR
                    KLFSFTPAWWTCKDLILLQALRESQSYLVEDLERS"

BASE COUNT       166 a       221 c      210 g      115 t
ORIGIN

Query Match     92.8%;   Score 687;  DB 9;  Length 712;
Best Local Similarity  100.0%;   Pred. No. 7.2e-99;
Matches 687;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    54 ATGGGGCGCGCGCGCGCGCCGCACCGCATCCTGGATGCCGCTGGAGAACCTGACCGCCGAGGAGCTC 113
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGGGCGCGCGCGCGCGCCGGACCGCCCATCCTGGATGCCGCTGGAGAAACCTGACCGCCGAGGAGCTC 60

Qy   114 AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCGCTGCGCGGAGGGGGCTACGGGCGCATCCCG 173
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCGCTGCGCGGAGGGGGCTACGGGCGCATCCCG 120

Qy   174 CGGGGCGCCGCCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTAC 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGGGGCGCCGGCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTAC 180

Qy   234 CTGGAGACCTACGGCGCCGAGCTCACCGCTAACGTTGCTCGCGACATGGGCCTGCAGGAG 293
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGGAGACCTACGGCGCCGAGCTCACCGCTAACGTGCTCGCGACATGGGCCTGCAGGAG 240

Qy   294 ATGGCCGGGGCAGCTGCAGGGGGCCAGCACCCAGCAGGGGGCTCTGGAGCCGGGCCCAGCTGGGATC 353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATGGCCGGGGCAGCTGCAGGGGGGCCAGCACGCACCCAGCAGGGGTCTGGAGCCGGGCCCAGCTGGGATC 300

Qy   354 CAGGCCCCTCCTCAGTCGGCGAGCCAAGCCAGGGCCGCTGCACTTTATAGACCAGCACCGGGGGCT 413
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGGCCCCTCCTCAGTCGGCGAGCCAAGGCCAGCGGCCGCTGCACTTTATAGACCAGCACCGGGGGCT 360

Qy   414 GCGCTTATCGCGAGGGTCACAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGGAAGGTC 473
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCGCTTATCGCGAGGGTCACAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGGAAGGTC 420

Qy   474 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCGGCCGAGCCCACCAACCCAAGCAAGATGCGG 533
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCGGCCGAGCCCACCAACCCAAGCAAGATGCGG 480

Qy   534 AAGCTCTTCAGTTTCACACCAGCCTGGGAACTTGGACCTGCAAAGGACTTGCTCCTCCAGGCC 593
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGCTCTTCAGTTTCACACCAGCCTGGGAACTTGGACCTGCAAAGGACTTGCTCCTCCAGGCC 540

Qy   594 CTAAGGGAGTCCCAGTCTTACCTGGTGGAGGACCTGGAGCGGGGGAGCTGAGGCTCCTTCCCA 653
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTAAGGGAGTCCCAGTCTTACCTGGTGGAGGACCTGGAGCGGGTGAGGCTCCTTCCCA 600

Qy   654 GCAACACTCCGGTCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGAATCTTTTTAT 713
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCAACACTCCGGTCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGAATCTTTTTAT 660

Qy   714 ACACAATATACGAAAAGCCAGCTTGAA 740
         ||||||||||||||||||||||||||||
Db   661 ACACAATATACGAAAAGCCAGCTTGAA 687
```

Search completed: January 29, 2004, 09:40:36
Job time : 2665.52 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.   GenCore version 5.1.6

Run on: January 29, 2004, 02:57:06 ; Search time 182.487 seconds
(without alignments)
10946.442 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcgggg.........atacgaaaagccagcttgaa 740

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 134971901? residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

```
 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	Description
1	740	100.0	740	22	AAF30007	Human CARD-5 cDNA.
2	740	100.0	740	24	ABK87966	Human caspase recr
c 3	740	100.0	740	24	ABK87967	Human caspase recr
4	740	100.0	740	25	ABS56032	cDNA encoding huma
5	734	99.2	770	22	AAD03890	Human target of me
6	734	99.2	782	24	AAL47126	Pyrin domain conta
7	733	99.1	745	22	AAS59817	Human novel cytoki
8	733	99.1	779	20	AAZ33631	Human breast tumou

No.	Score	%	Len	n	DB ID	Description
9	721.2	97.5	806	22	AAH34052	Human colon cancer
10	697.6	94.3	811	21	AAC77884	Human cancer assoc
11	610	82.4	713	21	AAD03906	Alternatively spli
12	531	71.8	639	22	AAH34053	Human colon cancer
13	384.6	52.0	432	21	AAC98638	Human colon cancer
14	378	51.1	405	22	AAD03907	Alternatively spli
15	360	48.6	2821	22	AAD03889	Human target of me
16	356	48.1	356	22	AAD03894	Human target of me
17	356	48.1	777	22	AAF30008	Mouse CARD-5 cDNA.
18	336	45.4	777	24	ABK87964	Mouse caspase recr
19	336	45.4	777	24	ABK87965	Mouse caspase recr
20	336	45.4	803	22	AAD03904	Mouse target of me
21	322	43.5	626	22	AAD03891	Human target of me
22	321	43.4	340	22	AAD03892	Human target of me
23	309.8	41.9	321	24	ABN94568	CpG island of huma
24	276.8	37.4	605	22	AAD03905	Rat target of meth
25	174.4	23.6	551	24	AAL47124	Pyrin domain conta
26	177.6	24.0	807	24	AAC80099	Human FLEXHT-30 nu
27	115	15.5	4194	22	AAD02761	Human NB-ARC and C
28	115	15.5	4329	22	AAD02762	Human NB-ARC and C
29	115	15.5	4422	22	AAD02760	Human NB-ARC and C
30	115	15.5	5100	24	AAL47127	Pyrin domain conta
31	115	15.5	5122	22	AAD12951	Human G-protein co
32	115	15.5	5523	22	AAF36651	Pyrin domain conta
33	115	15.5	5523	25	AAP83392	cDNA CARD-7 polyp
34	115	15.5	5523	25	AAS59497	cDNA encoding huma
35	111.4	15.1	2657	22	ABS56030	cDNA encoding huma
36	98.6	13.3	180	22	ABS55497	Human secreted pro
37	75.6	10.2	16891	20	AAX37084	Pyrin domain conta
38	74.2	10.0	3500	22	AAL47125	MEFV gene sequence
39	74.2	10.0	3512	20	AAX37085	Human novel cytoki
c 40	74	10.0	394	22	ABAS0929	Human breast cell
c 41	74	10.0	394	22	ABA68906	Human foetal liver
c 42	74	10.0	394	22	ABA35859	Probe #1425 for g
c 43	74	10.0	394	22	AAK17239	Human brain expres
c 44	74	10.0	394	22	AAK43026	Human bone marrow
c 45	74	10.0	394	22	AAI23791	Probe #13724 for g

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 1
AAF30007

```
ID   AAF30007 standard; cDNA; 740 BP.
XX
AC   AAF30007;
XX
DT   23-APR-2001  (first entry)
XX
DE   Human CARD-5 cDNA.
XX
KW   CARD-5; caspase recruitment domain; human; cancer; infection;
KW   autoimmune disease; neurological disease; haematological disease;
KW   immune disease; inflammation; antitumour; antiseptic;
KW   immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FH
FT   CDS            54..641
FT                  /*tag= a
FT                  /note= "the open reading frame is also specifically
FT                         claimed in Claim 1(a)"
XX
PN   WO200100826-A2.
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US17691.
XX
PR   28-JUN-1999; 99US-0340620.
XX
```

ALIGNMENTS

XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-061973/07.
DR P-PSDB; AAB20085.
XX
PT caspase activation are useful for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PT haematological disorders -
XX
PS Claim 1(a); Fig 21; 208pp; English.
XX
CC The present sequence is that of cDNA encoding human caspase
CC recruitment domain 5 (CARD-5, see AAB20085). The cDNA was isolated
CC from a testis cDNA library using murine CARD-L. Plasmid EpHC5
CC containing CARD-5 cDNA is deposited as ATCC PTA-213. CARD-5
CC is an intracellular protein predicted to be involved in regulating
CC caspase activation. It is useful as a modulating agent in regulating
CC regulating cellular processes include cell growth and cell death.
CC Methods of diagnosing and treating patients suffering from a
CC disorder associated with an abnormal level or rate of apoptotic
CC cell death, abnormal activity of the Fas/APO-1 receptor complex
CC abnormal activity of the tumour necrosis factor receptor complex
CC or abnormal activity of a caspase involve administering a compound
CC that modulates the expression or activity of CARD-3, CARD-4, CARD-5
CC or CARD-6 e.g. a small molecule, antisense nucleic acid, ribozyme
CC or polypeptide. Such disorders include cancer, viral infection,
CC autoimmune disorders, neurological diseases, haematological
CC disorders, inflammatory disorders and immune disorders. CARD
CC nucleic acids can be used to express CARD proteins in a host cell
CC e.g. for gene therapy applications, to detect a genetic lesion and
CC to modulate CARD activity.
XX
SQ Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match   100.0%; Score 740; DB 22; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.7e-147;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CGCGTCCGGCTGCAGCGGGGGTGAGCGGCGGCAGCGGGCCGGGGGATCCTGGAGCCATGGGGC     60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGCGTCCGGCTGCAGCGGGGGTGAGCGGCGGCAGCGGGCCGGGGGATCCTGGAGCCATGGGGC     60

Qy      61 GCGGGCGGCGACGCCATCCTTGGATCGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGT    120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCGGGCGGCGACGCCATCCTTGGATCGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGT    120

Qy     121 TCAAGCTGAAAGCTGCTGTCGGGTCGCGCTGCGCGAGCGGGCTACGGGCGCCATCCCGCGGGCG    180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCAAGCTGAAAGCTGCTGTCGGGTCGCGCTGCGCGAGCGGGCTACGGGCGCCATCCCGCGGGCG    180

Qy     181 CGCTGCGTGTCCATGGACGCCCTTGGACCTCACCGGACAAGCTGGTCAGCTTCTACCTGGAGA    240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CGCTGCGTGTCCATGGACGCCCTTGGACCTCACCGGACAAGCTGGTCAGCTTCTACCTGGAGA    240

Qy     241 CCTACGCGGCCCGAGCTCACCGGCTAACGTGTGCGCGACATGGGCGTGCAGGAGTGGCCG    300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTACGCGGCCCGAGCTCACCGGCTAACGTGTGCGCGACATGGGCGTGCAGGAGTGGCCG    300

Qy     301 GGCGACTGCAGGCGGCCACGCACCCAGGGCTCTGGAGCCCGCGCGGGATCCAGGGCC    360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGCGACTGCAGGCGGCCACGCACCCAGGGCTCTGGAGCCCGCGCGGGATCCAGGGCC    360

Qy     361 CTCCTCAGTGCGGCCAAGCCAGGCCTTGCACTTTATAGACACCACCCGGGCTGCGCTTA    420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTCCTCAGTGCGGCCAAGCCAGGCCTTGCACTTTATAGACACCACCCGGGCTGCGCTTA    420

Qy     421 TCGGGAGGGTCACAAACGTTGAGTGGCTGGARGCTCTGTACGGGAGGTCCTGACGG    480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCGGGAGGGTCACAAACGTTGAGTGGCTGGARGCTCTGTACGGGAGGTCCTGACGG    480

Qy     481 ATGAGCAGTACCAGGCAGTGCGGCGGGCGAGCCGCGAGCCCACCAAACCCAAGCAAGATCGCGGAAGCTCT    540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATGAGCAGTACCAGGCAGTGCGGCGGGCGAGCCGCGAGCCCACCAAACCCAAGCAAGATCGCGGAAGCTCT    540

Qy     541 TCAGTTTCACCACCAGCCTGGAACTGGACTGCAAGGACTTGCTCCTCCAGGCCCTAAGGG    600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCAGTTTCACCACCAGCCTGGAACTGGACTGCAAGGACTTGCTCCTCCAGGCCCTAAGGG    600

Qy     601 AGTCCCAGTCCTACCTGGGTGGAGGACCTGGAGGCGGCTCCTTCCCAGCAACAC    660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AGTCCCAGTCCTACCTGGGTGGAGGACCTGGAGGCGGCTCCTTCCCAGCAACAC    660

Qy     661 TCCGGTCAGCCCCTGGCAATCCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAAT    720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TCCGGTCAGCCCCTGGCAATCCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAAT    720

Qy     721 ATACGAAAAGCCAGCTTGAA    740
           ||||||||||||||||||||
Db     721 ATACGAAAAGCCAGCTTGAA    740
```

RESULT 2
ABK87966

ID ABK87966 standard; cDNA; 740 BP.
XX
AC ABK87966;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human caspase recruitment domain-5 (CARD-5) cDNA.
XX
KW Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;
KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
KW cell proliferation; gene therapy; immune disorder;
KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
KW stroke.
XX
OS Homo sapiens.
XX

FH Key   Location/Qualifiers
FT CDS   54..641
FT   /*tag= a
FT   /product= "CARD-5"
XX
PN WO200244354-A2.
XX
PD 06-JUN-2002.
XX
PF 29-NOV-2001; 2001WO-US044894.
XX
PR 01-DEC-2000; 2000US-0728721.
PR 24-APR-2001; 2001US-0841879.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2002-557538/59.
DR P-PSDB; AAU99353.
XX
PT Novel isolated murine or human caspase recruitment domain (CARD)-5
PT polypeptide, useful for treating immune disorders such as Hashimoto's
PT thyroiditis, graft rejection, allergy, glomerular nephritis,
PT tuberculosis -
XX

PS   Claim 7; Fig 3; 100pp; English.
XX
CC   The invention discloses the isolated polypeptides, and encoding nucleic
CC   acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
CC   (cysteinyl aspartate-specific proteinases) are central to the apoptotic
CC   program and responsible for the degradation of cellular proteins that
CC   lead to the morphological changes seen in cells undergoing apoptosis.
CC   Caspases interact with other caspases via their CARDs and different
CC   subtypes of CARDS may confer binding specificity. CARD-5 is an
CC   intracellular protein that is predicted to be involved in regulating
CC   caspase activation. CARD-5 activates the nuclear factor-kappa B
CC   (NF-kappaB) transcription factor pathway and binds the CARDs of
CC   caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
CC   activity and NF-kappaB activation, regulate cell growth and cell death
CC   and be used in gene therapy. The CARD-5 polypeptides are useful for
CC   identifying compounds which bind to them and modulate their activity and
CC   for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
CC   nucleic acids, antibodies and modulators of CARD-5 expression or activity
CC   can be used to treat immune disorders such as chronic inflammatory
CC   diseases and disorders, Hashimoto's thyroiditis, graft rejection,
CC   sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
CC   nephritis, human immunodeficiency virus (HIV) and bacterial infections
CC   (including tuberculosis and lepromatous leprosy) and in screening and
CC   detection assays. Modulators of CARD-5 activity or expression are also
CC   useful for treating autoimmune disorders, such as systemic lupus
CC   erythematosus and arthritis, cell depletion, neurological disorders,
CC   such as Alzheimer's disease, Parkinson's disease and spinal muscular
CC   atrophy, haematologic diseases, such as myelodysplastic syndrome and
CC   aplastic anaemia, myocardial infarction and stroke. The sequence
CC   presented is the human caspase recruitment domain-5 (CARD-5) cDNA.
XX
SQ   Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Qy   1   CGCGGTCCGGCGCTGCAGCGGGGGTGAGCGGCGCAGCGGCGCCGGGGATCCTTGAGCCCATGGGGC   60
Db   1   CGCGGTCCGGCGCTGCAGCGGGGGTGAGCGGCGCAGCGGCGCCGGGGATCCTTGAGCCCATGGGGC   60

Qy   61   GCGCGGCGCGCGCCGCCATCCTGGATGGGCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGT   120
Db   61   GCGCGGCGCGCGCCGCCATCCTGGATGGGCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGT   120

Qy   121   TCAAGCTGAAGCTGCTATCGGTGCCCGCTGCCCGAGGGCTACGGGGCCCATCCCGCGCGGGGCG   180
Db   121   TCAAGCTGAAGCTGCTGTCGGTGCCCGCTGCCCGAGGGCTACGGGGCCCATCCCGCGCGGGGCG   180

Qy   181   CGCTGCTGTCCATGGAACCTCACCGGACAGAAGTGGGTCAGCTTCTACCTGGAGA   240
Db   181   CGCTGCTGTCCATGGAACCTCACCGGACAGAAGTGGGTCAGCTTCTACCTGGAGA   240

Qy   241   CCTACGGCGCCGGAGTCACGGCTAACGTGTGTGCGGGCCTGCACATGGGCGTGCAGGAGATGGCCG   300
Db   241   CCTACGGCGCCGGAGTCACGGCTAACGTGTGTGCGGGCCTGCACATGGGCGTGCAGGAGATGGCCG   300

Qy   301   GGCAGCTGCAGGGGCGCCACCAGGGCGTTCTGGAGCCGCCCAGCTGGCATCCAGGCCC   360
Db   301   GGCAGCTGCAGGGGCGGCCACCAGGGCGTTCTGGAGCCGCCCAGCTGGCATCCAGGCCC   360

Qy   361   CTCCCTCAGTCGGCAGCCAAGCCAAGCCCACTTTATAGACCAGCACCGGGCTGGCGCTTA   420
Db   361   CTCCCTCAGTCGGCAGCCAAGCCAAGCCCACTTTATAGACCAGCACCGGGCTGCCGCTTA   420

Qy   421   TCGCGGAGGGGTCACAAAACGTTGAGTGGCCTGTCGTAACGGGAAGGTCCTGACCG   480
Db   421   TCGCGGAGGGGTCACAAAACGTTGAGTGGCCTGTCGTAACGGGAAGGTCCTGACCG   480

Qy   481   ATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGATGCGGGAAGCTCT   540
Db   481   ATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGATGCGGGAAGCTCT   540

Qy   541   TCAGTTTCACCACCAGCCTGGAACTGGAGCCTGCGAAGGAGGACTTGCTTCCTCCTCCAGGGGC   600
Db   541   TCAGTTTCACACCAGCCTGGAACTGGAGCCTGCGAAGGAGGACTTGCTTCCTCCTCCAGGGGC   600

Qy   601   AGTCCCAGTCCTACCTGGTGGAGGAGCTGAAGGCGTGGAGGCTCCTTCCTGCCAGCAACAC   660
Db   601   AGTCCCAGTCCTACCTGGTGGAGGAGCTGAAGGCGTGGAGGCTCCTTCCTGCCAGCAACAC   660

Qy   661   TCCGGTCAGCCCCTGGCCAATCCCACAAATCATCCTGGAATCTGGATCTTTTTTATACACAAT   720
Db   661   TCCGGTCAGCCCCTGGCCAATCCCACAAATCATCCTGGAATCTGGATCTTTTTTATACACAAT   720

Qy   721   ATACGAAAAGCCAGCTTGAA   740
Db   721   ATACGAAAAGCCAGCTTGAA   740

Query Match   100.0%;   Score 740;   DB 24;   Length 740;
Best Local Similarity   100.0%;   Pred. No.1;   1.7e-147;
Matches   740;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 3
ABK87967/c
ID   ABK87967 standard; DNA; 740 BP.
XX
AC   ABK87967;
XX
DT   07-OCT-2002   (first entry)
XX
DE   Human caspase recruitment domain-5 (CARD-5) cDNA complementary strand.
XX
OS   Homo sapiens.
XX
KW   myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
KW   stroke.
XX
DE   Human; ssj; caspase recruitment domain-5; CARD-5; antiinflammatory;
KW   immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
KW   apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
KW   cell proliferation; gene therapy; immune disorder;
KW   chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
KW   sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
KW   human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
KW   lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
KW   arthritis; cell depletion; neurological disorder; Alzheimer's disease;
KW   Parkinson's disease; spinal muscular atrophy; haematologic disease;
XX
CC   Novel isolated murine or human caspase recruitment domain (CARD)-5
CC   polypeptide, useful for treating immune disorders such as Hashimoto's
CC   thyroiditis, graft rejection, allergy, glomerular nephritis,
CC   tuberculosis
PN   WO20024354-A2.
XX
PD   06-JUN-2002.
XX
PF   29-NOV-2001; 2001WO-US44894.
XX
PR   01-DEC-2000; 2000US-0728721.
PR   24-APR-2001; 2001US-0841879.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI; 2002-557538/59.
XX
PT   Novel isolated murine or human caspase recruitment domain (CARD)-5
PT   polypeptide, useful for treating immune disorders such as Hashimoto's
PT   thyroiditis, graft rejection, allergy, glomerular nephritis,
PT   tuberculosis
XX
PS   Claim 8; Fig 3; 100pp; English.
XX
CC   The invention discloses the isolated polypeptides, and encoding nucleic
CC   acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
CC   (cysteinyl aspartate-specific proteinases) are central to the apoptotic
CC   program and responsible for the degradation of cellular proteins that
CC   lead to the morphological changes seen in cells undergoing apoptosis.
CC   Caspases interact with other caspases via their CARDS and different
CC   subtypes of CARDS may confer binding specificity. CARD-5 is an
CC   intracellular protein that is predicted to be involved in regulating
CC   caspase activation. CARD-5 activates the nuclear factor-kappa B

CC (NF-kappaB) transcription factor pathway and binds the CARDs of
CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
CC activity and NF-kappaB activation, regulate cell growth and cell death
CC and be used in gene therapy. The CARD-5 polypeptides are useful for
CC identifying compounds which bind to them and modulate their activity and
CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
CC can be used to treat immune disorders such as chronic inflammatory
CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
CC (including tuberculosis and lepromatous leprosy) and in screening and
CC detection assays. Modulators of CARD-5 activity or expression are also
CC useful for treating autoimmune disorders, such as systemic lupus
CC erythematosus and arthritis, cell depletion, neurological disorders,
CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
CC aplastic anaemia, myocardial infarction and stroke. The sequence
CC presented is the complementary strand to the human caspase recruitment
CC domain-5 (CARD-5) cDNA.
XX
SQ Sequence 740 BP; 120 A; 236 C; 238 G; 146 T; 0 other;

Query Match   100.0%; Score 740; DB 24;   Length 740;
Best Local Similarity   100.0%;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
                                  Pred. No. 1.7e-147;

```
Qy    1 CGCGTCCCGGCTGCAGCGGGCGGGGTGAGCGGCCCGCAGCGGCGGCGGGGGATCCTGGAGCCATGGGGC   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  740 CGCGGTCCGGCTGCAGCGGGGTGAGCGGCCCGCAGCGGCGGGGGATCCTGGAGCCATGGGGC  681

Qy   61 GGCGGCGGCGGACGCCATCCTGGATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  680 GGCGGCGGACGCCATCCTGGATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGT  621

Qy  121 TCAAGCTGAAGCTGCTGTCGGTCGGCCGCCGCTTACGGCGGCCATCCCCGCGGGGGCG  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  620 TCAAGCTGAAGCTGCTGTCGGTCGGCCGCCGCTTACGGCGGCCATCCCCGCGGGGGCG  561

Qy  181 CGCTTGCGTCCATGGACCCTTGGACTCACCGACAAGCTGGTTCAGCTTCTACCTGGAGA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  560 CGCTGCGTCCATGGACCGCTTCACCGACAAGCTGGTTCAGCTTCTACCTGGAGA  501

Qy  241 CCTACGGCGCCGAGCTCACCGCTAACGTGCTGCGGCCTGACAGGATGGCCG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  500 CCTACGGCGCCGAGCTCACCGCTAACGTGCTGCGGCCTGACAGGATGGCCG  441

Qy  301 GGCAGCTGCAGGCGGGCCACGCACCAGGGCGCTCGGAGCTCGGGATCCAGGGCCC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  440 GGCAGCTGCAGGCGGGCCACGCACCAGGGCGCTCGGAGCTCGGGATCCAGGGCCC  381

Qy  361 CTCCTCAGTCGGCGGCCAAGCCAGGCCTGCACTTTATAGGACCAGCAGCACGGGCCTTA  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  380 CTCCTCAGTCGGCGGCCAAGCCAGGCCTGCACTTTATAGGACCAGCAGCACGGGCTTA  321

Qy  421 TCGGCGAGGGTCACAAACGTTGAGTGGCTGCTGTAGGGGAGGTCCTGACGG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 TCGGGAGGGTCACAAACGTTGAGTGGCTGCTGTAGGGGAGGTCCTGACGG  261

Qy  481 ATGGAGCAGTACCAGGCAGTGCGGGCCGAGCCCACCAACCACCAAGATGCGGAAGCTCT  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  260 ATGGAGCAGTACCAGGCAGTGCGGGCCGAGCCCACCAACCACCAAGATGCGGAAGCTCT  201

Qy  541 TCAGTTTCACACAGGGGACAGGACTTGCTCCTCCAGGCCCTAAGGG  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  200 TCAGTTTCACACAGGGGAATGGACTTGCTCCTCCAGGCCCTAAGGG  141

Qy  601 AGTCCCAGTCCTACCTGGTGGAGGAGGCTCCTTCCCAGCAACAC  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  140 AGTCCCAGTCCTACCTGGTGGAGGAGGCTCCTTCCCAGCAACAC  81

Qy  661 TCCGGTCAGCCCCTGGCAATCCCACCCAAATCATCCTGAATCTGATCTTTTATACACAAT  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 TCCGGTCAGCCCCTGGCAATCCCACCCAAATCATCCTGAATCTGATCTTTTATACACAAT  21

Qy  721 ATACGAAAAGCCAGCTTGAA  740
        ||||||||||||||||||||
Db   20 ATACGAAAAGCCAGCTTGAA  1
```

RESULT 4

AC AB556032
ID AB556032 standard; cDNA; 740 BP.
XX
AC AB556032;
XX
DT 10-JAN-2003 (first entry)
XX
DE cDNA encoding human caspase recruitment domain-5 (CARD-5).
XX
KW Human; caspase activity; caspase recruitment domain-5; CARD-5;
KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
KW pseudoICE; ICEBERG; cell growth; cell death; inflammation; CARD-7;
KW apoptosis; caspase activation; cancer; follicular lymphoma;
KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
KW human immunodeficiency virus infection; HIV infection; ALS;
KW Alzheimer's disease; Parkinson's disease; amylotrophic lateral sclerosis;
KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
KW dermatological; nephrotropic; neuroprotective; cardiant; gene; ss;
KW caspase recruitment domain-7;
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   54..641
FT   /*tag= a
FT   /product= "CARD-5"
XX
PN US2002128198-A1.
XX
PD 12-SEP-2002.
XX
PF 27-NOV-2001; 2001US-0996617.
XX
PR 28-JUN-1999; 99US-0340620.
PR 27-OCT-1999; 99US-0428252.
PR 15-AUG-2001; 2001US-0931071.
XX
PA (BERT/) BERTIN J.
XX
PI Bertin J;
XX
DR WPI; 2003-028967/02.
DR N-PSDB; ABG71635.
XX
PT Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT CARD-7 and CARD-5 in presence of test compound, measuring their
PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PT altered.
XX
PS Disclosure; Fig 7; 43pp; English.
XX
CC The present invention relates to methods of identifying compounds
CC that regulate caspase activity using caspase recruitment domain-7
CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC method for identifying a compound that modulates the interaction
CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
CC CARD-8 molecules are useful as modulating agents in regulating a
CC variety of cellular processes including cell growth, cell death, and
CC inflammation. The methods of the invention are useful for identifying
CC compounds that have the ability to increase/decrease apoptosis, or

CC   comprise the ability to induce caspase activation. The methods are
CC   useful for treating a disorder associated with inappropriate apoptosis
CC   or inappropriate inflammation. The methods are useful for treating
CC   disorders associated with an undesirably low rate of apoptosis such
CC   as cancer (preferably follicular lymphoma, chronic myelogenous
CC   leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
CC   infections, autoimmune diseases caused by low levels of apoptosis
CC   (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC   and arthritis). The methods are also useful for treating disorders with
CC   undesirably high rates of apoptosis such as human immunodeficiency
CC   virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC   amyltrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC   muscular atrophy, various forms of cerebellar degeneration, anaemia
CC   associated with chronic disease, aplastic anaemia, chronic neutropaenia,
CC   myelodysplastic syndromes, myocardial infarction, stroke, and
CC   various inflammatory disorders (e.g. Crohn's disease, reactive
CC   arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC   psoriasis, graft rejection, allergic rhinitis, food allergies,
CC   conjunctivitis, glomerular nephritis, etc). The present sequence
CC   encoding human CARD-5 is used to identify cDNA encoding CARD-7.
XX
SQ   Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match   100.0%; Score 740; DB 25; Pred. No. 1.7e-147;
Best Local Similarity   100.0%;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CGGCGTCCGGCTGCAGCGGGGCGGGCAGCGGGCCCGGGGATCCTGAGCGCCATGGGGGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGGCGTCCGGCTGCAGCGGGGCGGGCAGCGGGCCCGGGGATCCTGAGCGCCATGGGGGC 60

Qy   61 GCGGCGGCCGACGCCATCCTTGGATGCCTGGGAGAACCTGACCGCCCGGGAGCTCAAGAAGT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCGGCGGCCGACGCCATCCTTGGATGCCTGGGAGAACCTGACCGCCCGGGAGCTCAAGAAGT 120

Qy   121 TCAAGCTGAAAGCTGCTGTGTGGGTGCCGCGAGGGCCATCCCGCGGGGGCG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCAAGCTGAAAGCTGCTGTGTGGGTGCCGCGAGGGCCATCCCGCGGGGGCG 180

Qy   181 CGCTGCTGCGTCCATGGACCGGCCTTTGGACCTCACCGACAAGCTGGGTCAGCTTCTTACCTGGAGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CGCTGCTGCGTCCATGGACCGGCCTTTGGACCTCACCGACAAGCTGGGTCAGCTTCTTACCTGGAGA 240

Qy   241 CCTACGGCGCCGGAGCTCACCGCTAACGTGCTGCGGCCGCTGACCATGGGCGTGGTCAGGAGGATGGCCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTACGGCGCCGGAGCTCACCGCTAACGTGCTGCGGCCGCTGACCATGGGCGTGGTCAGGAGGATGGCCG 300

Qy   301 GGCAGACTGCAGGGCGGCCCACGGCCACCAGGGGCTCTGGAGCCCGCGAGCTGGGGATCCAGGGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGCAGACTGCAGGGCGGCCCACGGCCACCAGGGGCTCTGGAGCCCGCGAGCTGGGGATCCAGGGCC 360

Qy   361 CTCCTCAGTCGGGCAGCCAAGCCAGGGCCTCCACTTTATAGACCAGCACCCGGGGCTTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTCCTCAGTCGGGCAGCCAAGCCAGGGCCTCCACTTTATAGACCAGCACCCGGGGCTTA 420

Qy   421 TCGCGGAGGGTCACAAACGTTGAGTGGTGCGGATGCTCTGTACGGAAGGTCCTGACGG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCGCGGAGGGTCACAAACGTTGAGTGGTGCGGATGCTCTGTACGGAAGGTCCTGACGG 480

Qy   481 ATGAGCAGTACCAGGCAGTGCGGGCCGACCCCACCAACCCAAGCAAGATCCGGAAGCTCT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCAGTACCAGGCAGTGCGGGCCGACCCCACCAACCCAAGCAAGATCCGGAAGCTCT 540

Qy   541 TCAGTTTCACACCAGCCTGGAAACTGGAACTGGCATTTGCACGCTTGCTCCTCCAGGCCCTTAAGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCAGTTTCACACCAGCCTGGAAACTGGAACTGGCATTTGCACGCTTGCTCCTCCAGGCCCTTAAGG 600

Qy   601 AGTCCCAGTTCACCTGGTTGGAGGAACCTGGAGGCTGAGGCTTCCTTCCCAGCCAACAC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGTCCCAGTTCACCTGGTTGGAGGAACCTGGAGGCTGAGGCTTCCTTCCCAGCCAACAC 660

Qy   661 TCCGGTTCAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTATACACAAT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCCGGTTCAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTATACACAAT 720

Qy   721 ATACGAAAAGCCAGCTTGAA 740
            ||||||||||||||||||||
Db   721 ATACGAAAAGCCAGCTTGAA 740

RESULT 5
AAD03890
ID   AAD03890 standard; cDNA; 770 BP.
XX
AC   AAD03890;
XX
DT   02-JUL-2001 (first entry)
XX
DE   Human target of methylation-induced silencing-1 (TMS1) cDNA.
XX
KW   Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   75..662
FT   /*tag= a
FT   /product= "Human TMS1 protein"
XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999; 99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI; 2001-290922/30.
DR   P-PSDB; AAB00588.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation useful for identifying subject at risk of developing tumor
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -
XX
PS   Claim 68; Page 113-114; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it.
CC   The present sequence is human target of methylation-induced silencing-1
CC   (TMS1) cDNA.
XX
SQ   Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

Query Match   99.2%; Score 734; DB 22; Pred. No. 3.2e-146;
Best Local Similarity   100.0%;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   7 CGGCTGCAGCGCGGGGGGTGAGCGCGGCGGGGGGGATCCTGGAGCGCCATGGGGGCGCGGC 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db     28  CGGCTGCAGCGGGGTGAGCGGGGGGATCCTGGAGCCATGGGCGGCGGC
           |||||||||||||||||||||||||||||||||||||||||||||||||
Qy     67  GCGGACGCCATCCCTGGATGGGGCTGGAGAACCTGGAGCTCAAGAAGTTCAAGC
           |||||||                           |||||||||||||||||||
Db     88  GGGACGCCATCCCTGGATGGGGCTGGAGAACCTGGAGCTCAAGAAGTTCAAGC     147
Qy    127  TGAAGCTGCTGTGTCGGTGCCGCTGGGAGGGGCTGGAGAAC...                186
           |||||||||||||||||||||||||||||||||||||||||
Db    148  TGAAGCTGCTGTGTCGGTGCCGCTGGGAGGGGCTGGAGAAC...                207
Qy    187  TGTCCATGGACGCCTTGGACCTCACCGACAAAGCTGGTCAGCTTCTTACCTGGAGGACCTACG     246
Db    208  TGTCCATGGACGCCTTGGACCTCACCGACAAAGCTGGTCAGCTTCTTACCTGGAGGACCTACG     267
Qy    247  GCGCCGAGCTCACCGCTAACGTGGGCCTGCAGGAGGATGGGCCGGGGCAGC     306
Db    268  GCGCCGAGCTCACCGCTAACGTGGGCCTGCAGGAGGATGGGCCGGGGCAGC     327
Qy    307  TGCAGGGCGGCCACGCACCAGGGCTCTGGAGACCCCGCCCAGCTGGGATCCCAGGGCCCCTCCTC     366
Db    328  TGCAGGGCGGCCACGCACCAGGGCTCTGGAGACCCGCCCAGCTGGGATCCCAGGGCCCCTCCTC     387
Qy    367  AGTCGGGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCACCGGGCGCTTATCGGCGA     426
Db    388  AGTCGGGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCACCGGGCGCTTATCGGCGA     447
Qy    427  GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTTCTGTACGGGAAGGTTCCTGACGGATGAGC     486
Db    448  GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTTCTGTACGGGAAGGTTCCTGACGGATGAGC     507
Qy    487  AGTACCAGGCCAGTGCGGCCGCCCACCAAACCCAAGCAAGATGCGGGAAGCTCTTCAGTT     546
Db    508  AGTACCAGGCCAGTGCGGCCGCCCACCAAACCCAAGCAAGATGCGGGAAGCTCTTCAGTT     567
Qy    547  TCACACCAGCCTGGAACTGGAACTTGCTCCTCCTCCAGGCCCTAAGGGAGTCCC     606
Db    568  TCACACCAGCCTGGAACTGGAACTTGCTCCTCCTCCAGGCCCTAAGGGAGTCCC     627
Qy    607  AGTCCTACTTGGTTGGAGCGAGCTGAGGCTGAGGCTCCCTCCTCCAGGCGCAACACTCCGGT     666
Db    628  AGTCCTACTTGGTTGGAGCGAGCTGAGGCTGAGGCTCCCTCCTCCAGGCGCAACACTCCGGT     687
Qy    667  CAGGCCCCTGGCAATCCACCAAAATCATCTCCTGAATCTGATCTTTTATACACAATATACGA     726
Db    688  CAGGCCCCTGGCAATCCACCAAAATCATCTCCTGAATCTGATCTTTTATACACAATATACGA     747
Qy    727  AAAGCCAGCTTGAA     740
Db    748  AAAGCCAGCTTGAA     761
```

RESULT 6

```
AAL47126
ID    AAL47126 standard; DNA; 782 BP.
XX
AC    AAL47126;
XX
DT    20-AUG-2002  (first entry)
XX
DE    Pyrin domain containing protein Pycard coding sequence.
XX
PD    23-MAY-2002.
XX
PF    30-OCT-2001; 2001WO-EP12545.
XX
PR    15-NOV-2000; 200DE-1056687.
PR    30-NOV-2000; 200DE-1059595.
XX
PA    (APOT-) APOTECH RES & DEV LTD.
XX
PI    Tschopp J, Martinon F;
XX
PS    Claim 5; Pig 1; 116pp; German.
XX
PN    WO200240668-A2.
PD    30-OCT-2001; 2001WO-EP12545.
```

New DNA encoding protein with pyrin domain, useful for treating
diseases involving impaired signal transduction, particularly
inflammation, also proteins and antibodies -

The present invention relates the DNA and their encoded proteins, where
the proteins contain at least one PYD (pyrin) domain. These can be used
to treat diseases associated with impaired intracellular signal
transduction, particularly inflammation such as psoriasis,
arteriosclerosis, bacterial or viral infections (particularly meningitis
and pneumonial, multiple sclerosis, rheumatoid arthritis, asthma,
sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
and Parkinson's diseases. The present sequence is a coding sequence of
the invention.

```
KW    Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW    antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW    neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW    nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW    inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW    arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW    osteoarthritis; glomerulonephritis; gene; dé.
OS    Unidentified.
```

```
DR    WPI; 2002-427093/45.
DR    P-PSDB; AAO17854.
```

```
Pyrin domain containing protein Pycard coding sequence.
```

```
SQ    Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other;

Query Match             99.2%;  Score 734;  DB 24;  Length 782;
Best Local Similarity   100.0%;  Pred. No. 3.2e-146;
Matches  734;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      7  CGGCTGCAGCGGGGGTGAGCGGGGGCAGCGGGGATCCTGGAGCCATGGGCGGCGGC     66
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40  CGGCTGCAGCGGGGGTGAGCGGGGGCAGCGGGGATCCTGGAGCCATGGGCGGCGGC     99
Qy     67  GCGACGCCATCCCTGGATGGGGCTGGAGAACCTGGAGCTCAAGAAGTTCAAGC     126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100  GCGACGCCATCCCTGGATGGGGCTGGAGAACCTGGAGCTCAAGAAGTTCAAGC     159
Qy    127  TGAAGCTGCTGTGTCGGTGCCGCTGGGAGGGGCTGGAGAACCTGACCGCGCGG     186
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160  TGAAGCTGCTGTGTCGGTGCCGCTGGGAGGGGCTGGAGAACCTGACCGCGCGG     219
Qy    187  TGTCCATGGACGCCTTGGACCTCACCGACAAAGCTGGTCAGCTTCTTACCTGGAGGACCTACG     246
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220  TGTCCATGGACGCCTTGGACCTCACCGACAAAGCTGGTCAGCTTCTTACCTGGAGGACCTACG     279
Qy    247  GCGCCGAGCTCACCGCTAACGTGGGCCTGCAGGAGGATGGGCCGGGGCAGC     306
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280  GCGCCGAGCTCACCGCTAACGTGGGCCTGCAGGAGGATGGGCCGGGGCAGC     339
Qy    307  TGCAGGGCGGCCACGCACCAGGGCTCTGGAGACCCCGCCCAGCTGGGATCCCAGGGCCCCTCCTC     366
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    340  TGCAGGGCGGCCACGCACCAGGGCTCTGGAGACCCCGCCCAGCTGGGATCCCAGGGCCCCTCCTC     399
Qy    367  AGTCGGGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCACCGGGCGCTTATCGGCGA     426
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400  AGTCGGGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCACCGGGCGCTTATCGGCGA     459
Qy    427  GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTTCTGTACGGGAAGGTTCCTGACGGATGAGC     486
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460  GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTTCTGTACGGGAAGGTTCCTGACGGATGAGC     519
Qy    487  AGTACCAGGCCAGTGCGGCCGCCCACCAAACCCAAGCAAGATGCGGGAAGCTCTTCAGTT     546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520  AGTACCAGGCCAGTGCGGCCGCCCACCAAACCCAAGCAAGATGCGGGAAGCTCTTCAGTT     579
```

```
Qy  547  TCACCAGCCTGGAACTGGACCTGCGAAGGACTTGCTCTCCTTCCAGGCGCCCTTPAGGGAGTCCC  606
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  580  TCACCAGCCTGGAACTGGACCTGCCAAGGACTTGCTCTCCTTCCAGGCGCCCTTAAGGGAGTCCC  639

Qy  607  AGTCCTACCTGGTGGAGGGAGGGAGGACTGAGGCTCCTTCCAGCGAACACTCCGGT  666
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  640  AGTCCTACCTGGTGGAGGGAGGGAGGACTGAGGCTCCTTCCAGCGAACACTCCGGT  699

Qy  667  CAGCCCCTGGCAATCCCACAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA  726
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  700  CAGCCCCTGGCAATCCCACAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA  759

Qy  727  AAAGGCCAGCTTGAA  740
         |||||||||||||||
Db  760  AAAGGCCAGCTTGAA  773
```

RESULT 7
AA559817

```
ID   AA559817 standard; cDNA; 745 BP.
XX
AC   AA559817;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human novel cytokine encoding cDNA 790CIP2B_1 #1.
XX
KW   Human; ss; cytokine; cell proliferation; cell differentiation;
KW   antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
KW   nervous system disease; neuropathy; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; spinal cord disorder;
KW   head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
KW   platelet disorder; thrombocytopaenia; stem cell disorder;
KW   aplastic anaemia; tissue regeneration; wound healing; ulcer;
KW   osteoporosis; osteoarthritis; bone degenerative disorder;
KW   periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
KW   severe combined immunodeficiency; infection; autoimmune disorder;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
KW   asthma; coagulation disorder; haemophilia; sepsis; nephritis;
KW   inflammatory bowel disease; food supplement.
XX
OS   Homo sapiens.
XX
XX
PN   WO200175093-A1.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US10484.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
PR   22-SEP-2000; 2000US-0668680.
PR   23-OCT-2000; 2000US-0695518.
PR   30-NOV-2000; 2000US-0728711.
PR   14-MAR-2001; 2000US-0728711.
XX
PA   (HYSB-) HYSEQ INC.
XX
PI   Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;
PI   Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;
XX
DR   WPI; 2001-626432/72.
DR   P-PSDB; AAU68525.
XX
PT   New polypeptides and nucleic acids, useful for diagnosis, treatment of
PT   inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
PT   degenerative disorders, cancer and promoting wound healing -
XX
PS   Claim 1; Page 150-151; 336pp; English.
XX
CC   The invention relates to isolated human polypeptides (which may be
CC   cytokines) and the polynucleotides encoding them. The protein is useful
CC   for identifying a compound which binds to it (e.g. modulators, agonists
CC   and antagonists). The polynucleotides are useful as an array for mismatch
CC   detection. The proteins and nucleic acids are useful as nutritional
CC   sources or supplements. The protein exhibits activity relating
CC   to cytokine, cell proliferation, cell differentiation, antiinflammatory,
CC   stem cell growth factor activity, immune stimulating or immune
CC   suppressing and activin or inhibin related activities. The proteins (and
CC   antibodies raised against them) and nucleic acids are therefore useful in
CC   the diagnosis and treatment of diseases and disorders such as cancer,
CC   central and peripheral nervous system diseases and neuropathies,
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
CC   diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
CC   thrombocytopaenia, stem cell disorders, aplastic anaemia, for
CC   regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC   growth, and in tissue repair, healing of burns, incisions, ulcers, for
CC   treating osteoporosis, osteoarthritis, bone degenerative disorders, or
CC   periodontal disease, lung or liver fibrosis, reperfusion injury in
CC   various tissues, various immune deficiencies and disorders including
CC   severe combined immunodeficiency (SCID), bacterial or fungal infections,
CC   autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
CC   diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
CC   such as asthma or other respiratory problems, coagulation disorders,
CC   haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory
CC   bowel disease, viral infection and are useful in altering bodily
CC   characteristics. The present sequence encodes a novel protein of the
CC   invention.
XX
SQ   Sequence 745 BP; 157 A; 234 C; 235 G; 119 T; 0 other;
```

```
Query Match      99.1%;  Score 733;  DB 22;  Length 745;
Best Local Similarity   100.0%;  Pred. No. 5.2e-146;
Matches 733;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy   8  GGCTCGCAGGCGGGCGGGTGAGCGCGGCACGCGGCGGGGGAATCCTGGAGCCCATGGGCGCGCGCG  67
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2  GGCTCGCAGGCGGGCGGGTGAGCGCGGCACGCGGCGGGGGAATCCTGGAGCCCATGGGCGCGCGCG  61

Qy  68  CGACGCCCATCCTGGATCGCGGCTGGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTTCAAGCT  127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  62  CGACGCCCATCCTGGATCGCGGCTGGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTTCAAGCT  121

Qy 128  GAAAGCTGCTGTCGGGTGCGGCGTGCCGCAGGGGCGCCATCCCGCGGGGGCGGCGCTGGCT  187
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122  GAAAGCTGCTGTCGGGTGCGGCGTGCCGCAGGGGCGCCATCCCGCGGGGGCGGCGCTGGCT  181

Qy 188  GTCCATGGACCGCCTTGGACCTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGG  247
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182  GTCCATGGACCGCCTTGGACCTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGG  241

Qy 248  CGCCGAGCTCACCGCTAAACGTTCGCCCGCAGGGAGCATGGGCCTGCAGGAGGAGATGGCCGGCAGCT  307
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 242  CGCCGAGCTCACCGCTAAACGTTCGCCCGACATGGGCCTGCAGGAGGAGATGGCCGGCAGCT  301

Qy 308  GCAGGGCGGCCACGCACCAGGGCTCTGGAGCCCGCCCAGCGTGGGATCCAGGGCCCCTCCTCA  367
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 302  GCAGGGCGGCCACGCACCAGGGCTCTGGAGCCCGCCCAGCGTGGGATCCAGGGCCCCTCCTCA  361

Qy 368  GTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACCACGCACCGGCTTATCGCGAG  427
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 362  GTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACCACGCACCGGCTTATCGCGAG  421

Qy 428  GGTCACAAACGTTGAGTGGCTGTGCGATGGCTTCGTACGGGAAGGTCCTGACGGGATGAGCA  487
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 422  GGTCACAAACGTTGAGTGGCTGTGCGATGGCTTCGTACGGGAAGGTCCTGACGGGATGAGCA  481

Qy 488  GTACCAGGCAGTGCGGGCCGAGCCCAACCCAAGCAAGATGCCGGAAGCTCTTCAGTTT  547
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 482  GTACCAGGCAGTGCGGGCCGAGCCCAACCCAAGCAAGATGCCGGAAGCTCTTCAGTTT  541

Qy 548  CACACCAGCCCTGGAACTGGACCTGCAAGGACTTGCTCTCCTTCCAGGCGCCCTTAAGGGAGTCCA  607
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 542  CACACCAGCCCTGGAACTGGACCTGCAAGGACTTGCTCTCCTTCCAGGCGCCCTTAAGGGAGTCCA  601
```

```
Qy         608 GTCCTACCTTGGTGAGGGAGGACTTGGAGGGAGCTGAGGCTTCCTTCCCAGCAACACACTCCCGGTC 667
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 GTCCTACCTTGGTGAGGGAGGACTTGGAGGGAGCTGAGGCTTCCTTCCCAGCAACACACTCCCGGTC 661

Qy         668 AGCCCCTGGCAATCCCACCAAAATCATCCTGAAATCATCTTTTTATACACAATATATACGAA 727
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 AGCCCCTGGCAATCCCACCAAAATCATCCTGAAATCATCTTTTTATACACAATATATACGAA 721

Qy         728 AAGCCAGCTTGAA 740
               |||||||||||||
Db         722 AAGCCAGCTTGAA 734
```

RESULT 8

```
AA233631
ID   AA233631 standard; cDNA; 779 BP.
XX
AC   AA233631;
XX
DT   08-DEC-1999 (first entry)
XX
DE   Human breast tumour-associated EST 21.
XX
KW   Expressed sequence tag; EST; human; breast; cancer; gene therapy;
KW   treatment; tumour; cytostatic; medicament; ss.
XX
OS   Homo sapiens.
XX
PN   DE19813839-A1.
XX
PD   23-SEP-1999.
XX
PF   20-MAR-1998; 98DE-1013839.
XX
PR   20-MAR-1998; 98DE-1013839.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;
XX
DR   WPI; 1999-528981/45.
XX
PT   Human nucleic acid sequences and protein products from tumor breast
PT   tissue, useful for breast cancer therapy -
XX
PS   Claim 1a; 101; 188pp; German.
XX
CC   This invention describes novel human nucleic acid sequences from tumor
CC   breast tissue which have cytostatic activity. The nucleic acid sequences
CC   can be used to produce and isolate full-length gene sequences. They can
CC   be used to express proteins, which can be used as tools to find an
CC   activity against breast cancer. The sequences can be used in sense or
CC   antisense form. They are especially useful for medicaments for gene
CC   therapy to treat breast cancer. AA233611-Z48617 represents expressed
CC   sequence tags described in the method of the invention.
XX
SQ   Sequence 779 BP; 149 A; 252 C; 251 G; 127 T; 0 other;

Query Match    99.1%; Score 733; DB 20; Length 779;
Best Local Similarity  100.0%;  Pred. No. 5.3e-146;
Matches  733; Conservative  0; Mismatches  0; Indels  0; Gaps  0;
```

```
Qy           7 CGGGTGCAGCGGGGGTGAGCGGGCGGCCGGGGGATCCTGGAGCCATGGGCCGCGCGCGC 66
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          47 CGGGTGCAGCGGGGGTGAGCGGGCGGCCGGGGGATCCTGGAGCCATGGGCCGCGCGCGC 106

Qy          67 GCGACGGCCATCCTGGATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTCAAGC 126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         107 GCGACGGCCATCCTGGATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTCAAGC 166

Qy         127 TGAAGCTGCTGTTCGGTTCGGTGCCGTCGGTGCCGCTACGGGGCGGGGGCATCCCGCGGGGGGGGCTGC 186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         167 TGAAGCTGCTGTTCGGTTCGGTGCCGTCGGTGCCGCTACGGGGCGGGGGCATCCCGCGGGGGGGGCTGC 226
```

```
Qy         187 TGTTCCATGGAGACGGCGCTTGGACCTCACCGACAGCAAGCTGGTCAGGCTTCAGCTTCTACCTGGAGACCTACG 246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         227 TGTTCCATGGAGACGGCGCTTGGACCTCACCGACAGCAAGCTGGTCAGGCTTCAGCTTCTACCTGGAGACCTACG 286

Qy         247 GCGCCGAGCTCACCGCCTAACGGTGTCGCGCCGAGCATGGGCGCCTGCAGCATGGGCCGCAGC 306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         287 GCGCCGAGCTCACCGCCTAACGGTGTCGCGCCGAGCATGGGCGCCTGCAGCATGGGCCGCAGC 346

Qy         307 TGCAGGCGGCGCCACCGACCCACGGCTTGGACCGGCGGCCGCAGCTTCCCTC 366
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         347 TGCAGGCGGCGCCACCGACCCACGGCTTGGACCGGCGGCCGCAGCTTCCCTC 406

Qy         367 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACCAGCACCGGGCGCTGCCGCTTATCGCGA 426
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         407 AGTCGGCAGCCAAGCCAGGGCCTGCACTTTATAGACCAGCACCGGGCGCTGCCGCTTATCGCGA 466

Qy         427 GGGTCACAAACGTTGAGTGGCTGCTGGAGTCGCTTCTCTGTACGGGAAGGTCCTGACGGATGAGC 486
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         467 GGGTCACAAACGTTGAGTGGCTGCTGGAGTCGCTTCTCTGTACGGGAAGGTCCTGACGGATGAGC 526

Qy         487 AGTACCAGGCAGTGCGGGCGCCGAGCCCACCAAACCCAAGCAAGATGCGGAAGCTTCTTCAGTT 546
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         527 AGTACCAGGCAGTGCGGGCGCCGAGCCCACCAAACCCAAGCAAGATGCGGAAGCTTCTTCAGTT 586

Qy         547 TCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCGCCCTAAGGGAGGAGTCCC 606
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         587 TCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCGCCCTAAGGGAGGAGTCCC 646

Qy         607 AGTCCTTACCTGGTGGAGCGGAGCGGAGCTGAGGCGTCAGGCGCAACAACCACTCCGGT 666
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         647 AGTCCTTACCTGGTGGAGCGGAGCGGAGCTGAGGCGTCAGGCGCAACAACCACTCCGGT 706

Qy         667 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTTGATCTTTTTATACACAATATACGA 726
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         707 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTTGATCTTTTTATACACAATATACGA 766

Qy         727 AAAGCCAGCTTGA 739
               |||||||||||||
Db         767 AAAGCCAGCTTGA 779
```

RESULT 9

```
AAH34052
ID   AAH34052 standard; cDNA; 806 BP.
XX
AC   AAH34052;
XX
DT   03-SEP-2001 (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:1134.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999; 99US-0157137.
PR   03-NOV-1999; 99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG74647.
XX
```

PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing; diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 2955; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAH77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 806 BP; 163 A; 253 C; 260 G; 128 T; 2 other;

Query Match   97.5%; Score 721.2; DB 22; Length 806;
Best Local Similarity   99.6%; Pred. No.1.7e-143;
Matches 731; Conservative   2; Mismatches   0; Indels   1; Gaps   1;

```
Qy           7 CGGCTGCACGGGGGTGAGGGGGGCAGCGGGGGGATCCTGGAGCCCATGGGGGGGCGCGCGC 66
               ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          68 CGGGCTGCACGGGGGTGAGCGGGCGCCAGCCCGGGGATCCTGGACCCATGGGGGGGCGCGCGC 127

Qy          67 GCCGACGCCATCCTGGATGCGCTGGAGAAACCTGACCTGAGGAGCTCAAGAAGTTCAAGC 126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         128 GCCGACGCCATCCTGGATGCGCTGCCGCGAGGGGCCATCCCGCGCGGGGGCGGGCTGC 187

Qy         127 TGAAGCTGCTGCTGGTCCGGTGCCCGTGCCGAGGGGCTACGGGGGCCATCCCGCGCGGGGGCGGGCTGC 186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         188 TGAAGCTGCTGGTCCGGTGCCCGGTGCCCGAGGGGCTACGGGGGCCATCCCGCGCGGGGGCGGGCTGC 247

Qy         187 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTGGTCAGCGTTCTACCTGGAGACCTACG 246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         248 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTGGTCAGCGTTCTACCTGGAGACCTACG 307

Qy         247 GCGCCGAGCTCACCGCTAACGTGTCTGCCGCGACCATGGGCTTCGAGGAAGATGGCCCGGGGCAGC 306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         308 GCGCCGAGCTCACCGCTAACGTGTCTGCCGCGACCATGGGCTTCGAGGAAGATGGCCCGGGGCAGC 367

Qy         307 TGCAGGCGGCCACCGCACCAGGGCTCTGGACCGGGGGGAATCCAGGGCCCCTCCTC 366
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         368 TGCAGGCGGCCACCGCACCAGGGCTCTGGACCGGGGGGGAATCCAGGGCCCCTCCTC 427

Qy         367 AGTCGGGCAGCCAAGCCCAGCCTGCACTTTATAGACCAGACACGCCGCTTATCGCCGA 426
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         428 AGTCGGGCAGCCAAGCCCAGCCTGCACTTTATAGACCAGACACGCCGCTTATCGCGA 487

Qy         427 GGGGTCACAAAGCGTTGAGTGGCTGCGTCCTGGGAATGCTCTCTGTACGGGAAGGTTCCTGACGGATGAGC 486
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         488 GGGGTCACAAAGCGTTGAGTGGCTGCGTCCTGGGAATGCTCTCTGTACGGGAAGGTTCCTGACGGATGAGC 547

Qy         487 AGTACCAGGCAGTGCGGGGCCGGAGCCCCACCAACCCCAAGGCAAGATGCCGGAGAAGCTCTTCAGTT 546
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         548 AGTACCAGGCAGTGCGGGGCCGGAGCCCCACCAACCCCAAGGCAAGATGGCGGAGAAGCTCTTCAGTT 607

Qy         547 TCACACCAGCCTGGAAACTGGAACTGGAACTTGCTCTCCTTCCCAGCGCCCTAAGGGAGTCCC 606
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         608 TCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTTCAGGCGCCCTAAGGGAGTCCC 667

Qy         607 AGTCCTACCTGGTTGGAGGAGGACCTGAGGAGGACTGGAACCTGAGGCTCCTTCCCAGCAACACTCCGGT 666
               |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db         667 AGTCCTACCTGGTTGGAGGAGGACCTGAGGACGGAGGCTCCTTCCCAGCAACACTCCGGT 666

Qy         667 CAGCCCCTGGCAAATCCCACCAAAATCATCCTGGAATTGATCTTTTTATACACAATATACGA 726
               || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         728 CA CCCCTGGCAAATCCCACCAAAATCATCCTGAATTCGATCTTTTTATACACAATATACGA 786

Qy         727 AAAGCCAGCTTGAA 740
               ||||||||||||||
Db         787 AAAGCCAGCTTGAA 800
```

RESULT 10
AAC77884
ID   AAC77884 standard; cDNA; 811 BP.
XX
AC   AAC77884;
XX
DT   08-FEB-2001 (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:278.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999; 99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43675.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reaction, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to

RESULT 11
XX
AC   AAC77884;
XX
DT   08-FEB-2001 (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:278.
XX
KW   neurological disease; drug screening; ss.
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
XX   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
XX   Human; cancer associated gene; cancer antigen; detection; cancer;
XX
Db   668 AGTCCTACCTGGTTGGAGGAGGACCTKGAGCGAGGAGGGACTGAGGAGGTCCTTCCCAGCAACACTCCGGT 727
               |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Qy   667 CAGCCCCTGGCAAATCCCACCAAAATCATCCTGGAATTGATCTTTTTATACACAATATACGA 726
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 CA-CCCCTGGCAAATCCCACCAAAATCATCCTGAATTCGATCTTTTTATACACAATATACGA 786
               || |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Qy   727 AAAGCCAGCTTGAA 740
               ||||||||||||||
Db   787 AAAGCCAGCTTGAA 800
```

```
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 811 BP; 170 A; 253 C; 259 G; 128 T; 1 other;

Query Match          94.3%;  Score 697.6;  DB 21;  Length 811;
Best Local Similarity 99.5%;  Pred. No. 1.6e-138;
Matches   730;  Conservative   1;  Mismatches   0;  Indels   3;  Gaps   3;

Qy      7 CGGCTGCAGCGGGGGGGTGAGCGGCGCAGCGGGCAGCCATGGGGGGCGCGCCC 66
          ||||||| |||||| |||||||||||||||||||||     ||||||||| ||||
Db     68 CGGCTGCGAGCGGGGGTGAGCGGCGCAGCGGGGATCCTGGAGCCATGGGGGCGC 127

Qy     67 GCGACGGCCATCCTGGATGCGGCTGGAGAACCTGACCCGCGCAGCGGGCCGCTGC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 GCGACGGCCATCCTGGATGCGGCTGGAGAACCTGACCCGCGCAGGAAGTTCAAGC 187

Qy    127 TGAAGCTGCTGTCGGTGCCGCTCGCCGGGGCCATCCCGCGGGGGGGGCCGCTGC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 TGAAGCTGCTGTCGGTGCCGCTCGCCGAGGGCTTACGGGGGCATCCCGCGGGGGG 247

Qy    187 TGTCCATGGACGGCCTTTGACCTCACCGGACAAGCTGGGTCAGCTTCTACCTGGA 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TGTCCATGGACGGCCTTTGGACCTCACCGGACAAGCTGGGTCAGCTTCTACCTGGA 307

Qy    247 GCGGCCGAGCTCACCGCTAACGTGCTGCGGGACCATGGGGCGCCGGAGGAGGCTCAAGAAGTTCAAGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GCGGCCGAGCTCACCGCTAACGTGCTGCGGGACCATGGGGCCTGCAGGAGGCTCAAGAAGTTCAAGC 367

Qy    307 TGCAGGGGGCGGCCACGCAGCCAGGGGTCTGGAGCCGGGATCCAGGGCCCCTCCTC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 TGCAGGGGGCGGCCACGCAGCCAGGGGTCTGGAGCCGGGATCCAGGGCCCCTCCTC 427

Qy    367 AGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACACCACCGCGCTGCGGCTTATTCGGGGA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 AGTCGGCAGCCAAGCCAGGCCTTGCACTTTATAGACACCACCGCGGCTTATTCGGGA 487

Qy    427 GGGTCACAAACGTTGACGTGGGCTGCTGCGATGCTCTGTACGGGGAAGGTCCTGACGGATGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 GGGTCACAAACGTTGAGTGGCTGCTGTACGGGGAAGGTCCTGACGGATGAGC 547

Qy    487 AGTACCAGGCAGTCGGGGCCGAGCCCACCAACCCAAGCAAGATGTCGGAAAGCTCTTCAGTT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 AGTACCAGGCAGTGC-GGCCGAGCCCACCAACCCAAGCAAGATGCGGAAAGCTCTTCAGTT 606

Qy    547 TCACACCAGCCTGGAACTGGACTTGCTCCTCTCCAGGGCCCTAAGGGAGTCCC 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    607 TCACACCAGCCTGGAACTGGACTTGCTCCTCTCCAGGGCCCTAAGGGAGTCCC 666

Qy    607 AGTCCTACCTTGGTGGGAGGACCTGGAGCGCTGAGCGGCGCTCCTTCCCAGCCAACACTCCGGT 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    667 AGTCCTACCTGGTGGGAGGACCTGGAGCC--GAGCTGAGGGCTGGGTCCTTCCCCACAACAACTCCGGT 725

Qy    667 CAGCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGAATCTTTTTTATACACAATATAGGA 726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 CA-SCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATTCTTTTTATACACAATATAGGA 784

Qy    727 AAAGCCAGCTTGAA 740
          ||||||||||||||
Db    785 AAAGCCAGCTTGAA 798
```

RESULT 12

AAH34053

ID   AAH34053 standard; cDNA; 639 BP.
XX
AC   AAH34053;
XX
DT   03-SEP-2001 (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:1135.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999; 99US-0157137.
PR   03-NOV-1999; 99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
PS   Claim 1; Page 2955; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with, inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated P,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 639 BP; 146 A; 198 C; 189 G; 106 T; 0 other;

Query Match   71.8%; Score 531; DB 22; Length 639;
Best Local Similarity   94.9%;
Matches 595; Conservative 0; Mismatches 20; Indels 12; Gaps 4;
                 Pred. No. 2.9e-103;

Qy   124 AGCTGAAGCTGTGCTTGCTGTCGTTCGGTCGCGCGAGGGGTCGGAGGGTCGGGCCATCCCGCG 176
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   2 AGCTGAAGCTGTGCTTGCTGTCGTTCGGTCGCGCGAGGGGTCGGAGGGTCGGGCCATCCTG 61

Qy   177 GGGCGCGGCTGCGTGTCCATGGACGCGCCTTGGGACCTTCACCGA----CAAGCTGGTCAGCTTCTAC 233
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   62 GATGCGCTGGGAGAAACCTGACCCGCCGAGGAGCTCAAGAAGTTCAAGCTGGTCAGCTTCTAC 121

Qy   234 CTGGAGACCTACGGGCGCGGGCGGAGCTCACGGCTAACGTGCTGCTGGGGGACCATGGGCCTGCAGGAG 293
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 CTGGAGACCTACGGGCGCGGGCGGAGCTCACCGCTAACGTGCTGCTGGGGGACCATGGGCCTGCAGGAG 181

Qy   294 ATGGGCCGGGCAGCTGCAGGCGCGGCCACGCCACCCAGGGGTCTCTGGAGCCGCGCCCAGCTGGGATC 353
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 ATGGGCCGGGCAGCTGCAGGCGCGGCCACGCCACCCAGGGGTCTCTGGAGCCGCGCCCAGCTGGGATC 241

Qy   354 CAGGCCCCTCCTCAGTCGCGCACCCAAGCCAGGCCAGGCCTCCACTTTATAGACCAGCACCGGGCT 413
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CAGGCCCCTCCTCAGTCGCGCAGCCAAGCCAGGCCAGGCCTCGCACTTTATAGACCAGCACCGGGCT 301

Qy   414 GCGCTTATCGCCGAGGGGTCACAAAACGTTGAGTGGCGCTCGTCGTTACGGGGAAGGTC 473
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GCGCTTATCGCCGAGGGGTCACAAAACGTTGAGTGGCGCTCGTCGTTACGGGGAAGGTC 361

Qy   474 CTGACCGGATGAGACAGTACCAGGCAGTAGCCGCGGAGCCCACCCAACCCAAGCAAGATCGGG 533
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 CTGACCGGATGAGACAGTACCAGGCAGTGC-GGCGGAGCCCACCCAACCCAAGAGATCGGG 420

Qy   534 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGGCAAGGACTTGCTCCTCCCAGGCC 593
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGGCAAGGACTTGCTCCTCCCAGGCC 480

Qy   594 CTAAGGGAGTCCCAGTCCTACTTGGTGGGAGGAACCTGGAGCGAGGCTGAGGGTCCTCCA 653
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTAAGGGAGTCCCAGTCCTACTTGGTGGGAGGACCTGGAGCGAGGCTGAGGGTCCTTCCA 540

Qy   654 GCAACACTTCCGGTCAGCGCCCCTTGGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTAT 713
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCAACACTTCCGGTCAG--CCCTGGCAACCCCACCAAAATCATCCTGAATCTGATCTTTTTAT 599

Qy   714 ACACAATATTACGAAAAAGCCAGCGTTGAA 740
            |||||||||||||||||||||||||||||
Db   600 ACACAATATTACGAAAAAGCCAGCGTTGAA 626

RESULT 13

AAC98638

ID   AAC98638 standard; cDNA; 432 BP.
XX

Db   268 GCGCCGGAGCTCACCGGTAACGTGCTCGCGGATGGGCGCTGCAGGAGGATGGCCGCGGGCCAGC 327
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   307 TGCAGGCGGCCACGCACCAGGGGTCTCTGGAGCCGCGCCCAGCTGGGATCCAGGCCCCTCCTC 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 TGCAGGCGGCCACGCACCAG-------------------------------- 347

Qy   367 AGTCGCGCAAGCCAGGCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTATCGCCGA 426
            ||||||||||||||||||||||||||||||||||||||||||
Db   348 --------------GGCTGCACTTTATAGACCAGCACCGGGCTGCGCTTATCGGCGA 390

Qy   427 GGGGTCACAAAACGTTGAGTGGCGCTCGTCGTTACGGGGAAGGTCCTGACCGGATGAGAGC 486
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   391 GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGACGGATGAGC 450

Qy   487 AGTACCAGGCAGTGCGGCGGCGGAGCCCACCCAACCCAAGCAAGATGCGGAAGCTTCTTCAGTT 546
            |||||||||||||||||||||||||||||||||||||||||
Db   451 AGTACCAGGCAGTGCGGCGGCCGGAGCCCACCCAAGCAAGAGTCGGAAGGCTCTTCAGTT 510

Qy   511 TCACACCAGCCTGGGAACTGGACTTGCTCCTCCCAGGGCCCTAAGGGAGTTCCC 570
            ||||||||||||||||||||||||||||||||||||||||||
Db   547 TCACACCAGCCTGGGAACTGGACTTGGCTGCAGGGGACTCCC 606

Qy   607 AGTCTCTACCTGGTGGAGGAACCTGGAGCGGGGAGGACTTGTCCTCCTCCCAGACACACTCCGGT 666
            ||||||||||||||||||||||||||||||||||||||||||
Db   571 AGTCCTACCTGGTGGAGGAACCTGGAGCGAGGCTGAGGGTCCTGACGAACACACTCCGGT 630

Qy   667 CAGCCCCTGGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTATACACAATATACGA 726
            ||||||||||||||||||||||||||||||||||||||||||
Db   667 CAGCCCCTGGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTTATACACAATATTACGA 690

Qy   631 CAGCCCCTGGGCAATCCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTTATACACAATATACGA 690

Qy   727 AAAGCCAGCGTTGAA 740
            |||||||||||||
Db   691 AAAGCCAGCTTGAA 704

Qy   547 TCACACCAGCCTGGGACTGGACCTGCAGCACCCAGGGCCCTTAAGGGAGTCCC 606
Qy   451 AGTACCAGGCAGTGGGGCCGGAGCCCACCCAAGCAAGATGGGGAAGGTTCTTCAGTT 510
Qy   391 GGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGACGGATGAGC 450
Qy   348 ------------------GGCTGCACTTTATAGACCAGCACCGGGCTTATGCGGCGA 390
Qy   328 TGCAGGCGGCCACGCACCAG-------------------------------- 347
Qy   268 GCGCCGGAGCTCACCGGTAACGTGCTCGCGGATGGGCGCTGCAGGAGGATGGCCGCGGGCCAGC 327

```
AC   AAC98638;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen nucleotide sequence SEQ ID NO:648.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder; ss.
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05883.
XX
PR   12-MAR-1999; 99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   P-PSDB; AAB53881.
XX
PS   Claim 1; Page 1175; 2104pp; English.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer -
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular,
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins
CC   may also be used to prevent diseases such as neural disorders, immune
CC   system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, wounds, renal disorders, infectious
CC   diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC   AAB54007 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 432 BP; 77 A; 142 C; 152 G; 61 T; 0 other;

Query Match             52.0%;  Score 384.6;  DB 21;  Length 432;
Best Local Similarity   98.4%;
Matches  420;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy      9 GCTGCAGCGGGGGTCAGCGCGGCAGCGCGGCCGCCGGGGATCCTGGAGCCGCGGCGCGCGC 68
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      8 GCTGCAGCGGGGGTGAGCGCGGCAGCGCGGCCGCCGGGGATCCTGGAGCCGCGGCGCGCGC 67

Qy     69 GACGGCCATCCTGGATGCGCGCGGAGAACCTGACCCGCGGAGAGCTCAAGAAGTTCAAGCTG 128
          |||||||||||||||||||| ||||||| |||||||||| |||||||||||||||||||||
Db     68 GACGGCCATCCTGGATGCGCGGAGAAACCTGACCCGCGAGGAGCTCAAGAAGTTCAAGCTG 127

Qy    129 AAGCTGCTGTCGGTGCGCGGTGCCCGCTGCCCGAGGCCGCTACGGCGCGGGCGCGGCTGCTG 188
          ||||||||||||||||||| ||||||||||| |||||| || ||| |||||||||||||||
Db    128 AAGCTGCTGTCGGTGCGCGGCGCGCCTGCCCGCGAGGCCATCCCGCCGGGGCGCGGCTGCTG 187

Qy    189 TCCATGGAGGAGCCCTTGGACCTCACCGACCAAGCTGTGTCAGCTTCTACCTGGAGACCTACGGC 248
```

```
Db    188 TCCATGGACGGCCTTGGACCTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGGC 247
          |||||||  |||| |||||||||||||||||||||  ||||||||||||||||||||||||||
Qy    249 GCCGAGCTTCACCGCTAAGGTGCTTGCGCGCCACACGACAAGCTGGTTCAGCGAGAGAGATGGCCGGGCGGCAGCTG 308
          |||||||| |||| ||||||||||| |||||||||||||  ||||||  ||| |||||| ||||  ||||||||||||
Db    248 GCCGAGCTCACCCGTAAGGTGCTGCGCGCCACACGATGGGCCTCCAGGAGAGATGGCGGGGCGGCAGCTG 307

Qy    309 CAGGCGGGGCCAGCGACCACCAGGCGCTTGGAGGCCGCGGCTGGGGATCCGGGATCCAGGGCCTCCTCAG 368
          ||||||| || ||| |||||| |||| |||||||||||||| ||||||||||| |||||||||||||||
Db    309 CAGGCGGCGCACCGACCATGGGCCTTGGGGATCCAGGGCCTCCTCAG 367

Qy    369 TCGGCAGCCAAGCCCAAGGCCTGCACTTT-ATAGACCAGCACCGGGCTGCCCTTATCGCGAG 427
          |||||||||||||||||| || ||||  | ||||||||||||||||| | |  ||||||| |
Db    367 TCGGCAGCCAAGCCCAAGCCCGCACCTTAATAGACCAGCACCGGGCTTCG-TTATCGCGAA 425

Qy    428 GGTCACA 434
          |||||
Db    426 GGTCAAA 432
```

```
RESULT 14
AAD03907
ID   AAD03907 standard; cDNA; 405 BP.
XX
AC   AAD03907;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Alternatively spliced form of human TMS1 cDNA (lacking exon3).
XX
KW   Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast; ss.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             75..404
FT                   /tag=
FT                   /product= a
FT                   /note= "Alternatively spliced form of human TMS1
FT                   protein lacking exon3"
FT                   /partial
FT                   "CDS does not include stop codon"
XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999; 99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI; 2001-290922/30.
DR   P-PSDB; AAE00595.
XX
PS   Claim 106; Page 123-124; 124pp; English.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation useful for identifying subject at risk of developing tumor
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
```

CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it.
CC   The present sequence is alternatively spliced form of human target of
CC   methylation-induced silencing-1 (TMS1) cDNA lacking exon3.
CC   Note: The present sequence designated as SEQ ID NO:26, is referred
CC   as DNA throughout the specification, however this sequence has been
CC   referred as protein in claim 106.
XX
SQ   Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other;

Query Match   51.1%; Score 378; DB 22; Length 405;
Best Local Similarity   100.0%; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
Matches 378; Pred. No. 6.6e-71;

Qy   7 CGGCCTGCACGGGGGGTGAGCCGCGGGGGATCCTTGGACGCCCATGGGGGGGCGGCGCGGCGC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28 CGGCCTGCACGGGGGGTGAGCCGCGGGGGATCCTTGGACGCCCATGGGGGGGCGGCGCGGCGC 87

Qy   67 GCGACGCCATCCTGGATGCGCGGGAGAAACCTGACCGCCGCGGGAGTCAAGAAGTTCAAGC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   88 GCGACGCCATCCTGGATGCGCGGGAGAAACCTGACCGCCGCGGGAGTCAAGAAGTTCAAGC 147

Qy   127 TGAAGCTGCTGTGTCGGTGCCCCGCGGGGCTACGGGCCATCCCGCGGGGGGCGCGGCGGCTGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148 TGAAGCTGCTGTGTCGGTGCCCCGCGGGGCTACGGGCCATCCCGCGGGGGGCGCGGCGGCTGC 207

Qy   187 TGTCCATGGACGCGCCTTGGACTCACGACAAGCTCGGTCAGCTTCTACCTTGGAGACCTACG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 TGTCCATGGACGCGCCTTGGACTCACGACAAGCTCGGTCAGCTTCTACCTTGGAGACCTACG 267

Qy   247 GCGCGCGAGCTCACCGCTAACGTGCTGCGCGACATGGGGCCTTGCAGGAGATGGCCGGGCGCAGC 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 GCGCGCGAGCTCACCGCTAACGTGCTGCGCGACATGGGGCCTTGCAGGAGATGGCCGGGCGCAGC 327

Qy   307 TGCAGGGCGGCCACCGCCGCTCTGGACCGTGGGGATTCCAGGGCCCTCCTC 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 TGCAGGGCGGCCACCGCCGCTCTGGACCGTGGGGATTCCAGGGCCCTCCTC 387

Qy   367 AGTCCGGCGCAGCCAAGCCAG 384
              |||||||||||||||||||||
Db   388 AGTCCGGCGCAGCCAAGCCAG 405

RESULT 15
AAD03889
ID   AAD03889 standard; DNA; 2821 BP.
XX
AC   AAD03889;
XX
DT   02-JUL-2001 (first entry)
XX
DE   Human target of methylation-induced silencing-1 (TMS1) genomic DNA.
XX
KW   Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast; ds.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   misc_feature   1100..1725
FT   /tag= a
FT   /label= CpG island
FT   /note= "this region is specifically claimed in claim 71"
FT   CDS   1256..2361
FT   /*tag= b
FT   /product= "CDS is interrupted by 2 introns"
FT   exon   1256..1529

FT   intron   1530..1742
FT   /*tag= c
FT   /number= 1
FT   exon   1743..1799
FT   /*tag= d
FT   /number= 2
FT   intron   1800..2104
FT   /*tag= e
FT   /number= 2
FT   exon   2105..2361
FT   /*tag= f
FT   /number= 3

XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999; 99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI; 2001-290922/30.
DR   P-PSDB; AAE00588.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation useful for identifying subject at risk of developing tumor
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -
XX
PS   Claim 68; Page 112-113; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it.
CC   The present sequence is human target of methylation-induced silencing-1
CC   (TMS1) genomic DNA.
XX
SQ   Sequence 2821 BP; 614 A; 864 C; 757 G; 586 T; 0 other;

Query Match   48.6%; Score 360; DB 22; Length 2821;
Best Local Similarity   100.0%; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
Matches 360; Pred. No. 5.6e-67;

Qy   381 CCAGGCCTGCACTTTATAGACCAGCACCGGGGCTGCCCTTATCGGACGGGTCACAAAACGTT 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CCAGGCCTGCACTTTATAGACCAGCACCGGGGCTGCCCTTATCGGACGGGTCACAAAACGTT 2160

Qy   441 GAGTGGCTGCTGGATGCTCTGTACGCGGAAGGTCCTGACGGGATGAGCAGTACCAGGCAGTG 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GAGTGGCTGCTGGATGCTCTGTACGCGGAAGGTCCTGACGGGATGAGCAGTACCAGGCAGTG 2220

Qy   501 CGGGGCCGAGCCCACCAACCCAAGCCAACAGATGCGGAAGCTTCTTCAGTTCACCAGCCTGG 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CGGGGCCGAGCCCACCAACCCAAGCCAACAGATGCGGAAGCTTCTTCAGTTCACCAGCCTGG 2280

Qy   561 AACTGGACCTGCAAGGACTTGCTTCCTCCCCAGGAGCCCTAAGGGAGTCCCAGTCCTACCTGGTG 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db  2281 AACTGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGGGAGTCCCAGTCCTACCTGGTG 2340

Qy   621 GAGGACCTGGGAGCCGGAGCTGAGAGCTCCTTCCCAGCCAACACTCCGGGTCAGCCCTGGCAAT  680
         ||||||||||||||||||||||||||||| |||||||||||||||||| |||||||||||||||
Db  2341 GAGGACCTGGGAGCCGGAGCTGAGGGCTCCTTCCCAGCCAACACTCCGGTCAGCCCTGGCAAT 2400

Qy   681 CCCACCAAAATCAATCCTGAATCTGAATCTTTTTATACACAATATACGAAAAGCCAGCTTGAA  740
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CCCACCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGAAAAGCCAGCTTGAA 2460
```

Search completed: January 29, 2004, 03:29:23
Job time : 184.487 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 03:01:52 ; Search time 1474.73 Seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-7

Perfect score: 740

Sequence: 1 cgcgtccggctgcagcggggg..........atacgaaaagccagcttgaa 740

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 121522238056 residues

Total number of hits satisfying chosen parameters: 45562784

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :

	EST:*		
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estmu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_htc:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_htc:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estun:*		
16:	em_estom:*		
17:	em_gss_hum:*		
18:	em_gss_inv:*		
19:	em_gss_pln:*		
20:	em_gss_vrt:*		
21:	em_gss_fun:*		
22:	em_gss_mam:*		
23:	em_gss_mus:*		
24:	em_gss_pro:*		
25:	em_gss_rod:*		
26:	em_gss_phg:*		
27:	em_gss_vrl:*		
28:	gb_gss1:*		
29:	gb_gss2:*		

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	729.4	98.6	930	12	BM459678	BM459678 AGENCOURT
2	727	98.2	804	12	BM926438	BM926438 AGENCOURT
3	722	97.6	767	12	BM910918	BM910918 AGENCOURT
4	721.2	97.5	792	12	BM549665	BM549665 AGENCOURT

ALIGNMENTS

RESULT 1
BM459678
BM459678

LOCUS   BM459678   930 bp   mRNA   linear   EST 05-FEB-2002
DEFINITION AGENCOURT_6417789 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534308
            5', mRNA sequence.
ACCESSION   BM459678
VERSION   BM459678.1 GI:18508718
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 930)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12220 row: j column: 05
            High quality sequence stop: 670.

c	5	713.4	96.4	750	12	BG764161	BG764161 602737190
c	6	704.8	95.2	725	14	CA309883	CA309883 UI-H-FTI-
c	7	690.8	93.4	856	12	BG767422	BG767422 602741219
c	8	688.6	93.1	748	10	BG337806	BG337806 602435709
c	9	670.4	90.6	678	9	AI148558	AI148558 qc68g10.x
c	10	667.6	90.2	741	12	BM051141	BM051141 603634063
c	11	660	89.2	676	12	BM972785	BM972785 UI-E-EJ0-
c	12	654.2	88.4	675	10	BE908204	BE908204 601500461
c	13	647.4	87.5	744	10	BG255521	BG255521 602367671
c	14	639	86.4	687	10	BE902218	BE902218 601501478
c	15	617.2	83.4	638	12	BM998354	BM998354 UI-H-DT1-
c	16	610	82.4	1079	12	BM456838	BM456838 UI-E-EJ0-
c	17	606	81.9	633	12	BM681310	BM681310 UI-E-EJ0-
c	18	604.6	81.7	1201	13	BX356773	BX356773 UI-E-EJ0-
c	19	581.2	78.5	830	10	BF238010	BF238010 601811713
c	20	553.8	74.8	621	12	BI462049	BI462049 UI-E-EJ0-
c	21	546	73.8	866	10	BG684386	BG684386 602636137
c	22	544.6	73.6	634	9	AI991236	AI991236 wu41c09.x
c	23	539	72.8	878	13	BU501178	BU501178 AGENCOURT
c	24	536	72.4	888	10	BE906296	BE906296 601502291
c	25	533	72.0	615	9	AI346818	AI346818 qp58f05.x
c	26	511.4	69.1	542	12	BQ045282	BQ045282 UI-CF-EN1
c	27	488.4	66.0	518	13	BU626120	BU626120 UI-H-FG1-
c	28	477.4	64.5	567	9	AA528254	AA528254 nh92d02.s
c	29	471	63.6	572	14	CA439478	CA439478 UI-H-D10-
c	30	469.2	63.4	520	12	BI468250	BI468250 id86e08.y
c	31	468.4	63.3	486	12	BQ012745	BQ012745 UI-1-BC1p
c	32	463	62.6	500	10	AW973297	AW973297 EST385395
c	33	450.4	60.9	481	9	AA573948	AA573948 nk09f05.s
c	34	446	60.3	446	12	BM726665	BM726665 UI-E-EJ0-
c	35	441.8	59.7	465	12	AA582937	AA582937 nn80b06.s
c	36	431.6	58.3	700	14	CB429243	CB429243 605047 MA
c	37	430.4	58.2	698	14	CB452677	CB452677 707559 MA
c	38	426	57.6	695	14	CB423226	CB423226 595419 MA
c	39	420.6	56.8	1093	12	BM920038	BM920038 AGENCOURT
c	40	409.8	55.4	421	9	AA278825	AA278825 zs78f06.s
c	41	408	55.1	671	14	CB421971	CB421971 595035 MA
c	42	407	55.0	669	14	CB429601	CB429601 605431 MA
c	43	390	52.7	565	14	BM362328	BM362328 BS320043
c	44	385.4	52.1	583	14	CB436146	CB436146 616467 MA
c	45	382	51.6	395	9	AI262374	AI262374 qk37a01.x

RESULT 2

FEATURES   Location/Qualifiers
  source   1..930
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5534308"
                  /tissue_type="leiomyosarcoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_71"
                  /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT."

BASE COUNT   182 a   310 c   263 g   172 t   3 others
ORIGIN

Query Match   98.6%;   Score 729.4;   DB 12;   Length 930;
Best Local Similarity   99.7%;   Pred. No. 8.5e-154;
Matches 730; Conservative   0; Mismatches   2;   Indels   0;   Gaps   0;

Qy   9 GCTGCGCAGCGGGGTGAGCGGCGGCGGAGCGGCGGCGGGGGAGATCCTGAGCCATGGGGCGGGGCGGGGCGGC 68
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCTGCGCAGCGGGGTGAGCGGCGGCGGAGCGGCGGCGGGGGAGATCCTGAGCCATGGGGCGGGGCGGGGCGGC 60

Qy   69 GACGCCACCCTGGATGCGCGTGGAGAACCTGACCGCCCCGGGGGAGCGCCAAGAAGTTCAAGCTG 128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GACGCCCATCCTGGATGCGCGTGGAGAACCTGACCGCCCGGGGGAGCGCCAAGAAGTTCAAGCTG 120

Qy   129 AAGCTGCTGTCGGGTGCGCTGCGCGGAGGGCTACGGCGGGGGGGCGGGGGCGGGGGCGGCCTGCTG 188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAGCTGCTGTCGGGTGCGCTGCGCGGAGGGCTACGGCGGGGGGGCGGGGGCGGGGGCGGCCTGCTG 180

Qy   189 TCCATGGACGCCTTGGACTTCACCGACAAAGCTGGTCAGCTTCTACCTGGAGACCTACGGC 248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TCCATGGACGCCTTGGACTTCACCGACAAAGCTGGTCAGCTTCTACCTGGAGACCTACGGC 240

Qy   249 GCCGAGCTCACCGCTAACTTGCTGCGGGCCTCAGAGAGATGGGCCGGGGCAGCTG 308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCCGAGCTCACCGCTAACGTGCTGCGGGCCTCAGAGAGATGGGCCGGGGCAGCTG 300

Qy   309 CAGCGCCAGCTCAGGCCTTGGGAGCCGCGCGCCCAGCTGGGGATCCAGGCGCCCTTCCTCAG 368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGCGCCAGCACCAGGGCTCTGGGAGCCGCGCGCCCAGCTGGGGATCCAGGCGCCCTTCCTCAG 360

Qy   369 TCGGCGCAGCCAAGCCAGGCCTGCCACTTTATAGACCAGCACCGGGCTGCGCTTATCGCGAGG 428
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCGGCGCAGCCAAGCCAGGCCTGCCACTTTATAGACCAGCACCGGGCTGCGCTTATCGCGAGG 420

Qy   429 GTCACAAAGCTTGAGCTGCTGGATGCTCTGTACGGAAGGTGCCTGACGGATGAGCAG 488
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTCACAAACGTTGAGCTGGCTGCTGGATGCTCTGTACGGGAAGGTGCCTGACGGATGAGCAG 480

Qy   489 TACCAGGCCAGTGCGGGCCACCAACCCAAGCAAGATGCGGAAGCTGTCTTCAGTTTC 548
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TACCAGGGCAGTGCGGGCCCAGCCCACCAAGCAAGATGCGGAAGCTGTCTTCAGTTTC 540

Qy   549 ACACCAGCCTGGAACTGGAACCTGCTGTCCTCCTCCCAGGCCCTAAGGGAGGTCCCAG 608
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACACCAGCCTGGAACTGGAACCTGCTGTCCTCCTCCCAGGCCCTAAGGGAGGTCCCAG 600

Qy   609 TCCTTACCTGGTGGAGGAGGACTGGAGGCTCTTCTTCCAGCCACACACTCCGGTCA 668
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCCTTACCTGGTGGAGGAGGACTGGAGGCTCTTCTTCCAGCCACACACTCCGGTCA 660

Qy   669 GCCCCTGGCAATCCCACAAAATCATCCTGGAATCTGAATCTTTTTATACACAATATACGAAA 728
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GCCCCTGGCAATCCCACAAAATCATCCTGGAATCTGAATCTTTTTATACACAATATACGAAA 720

Qy   729 AGCCAGCTTGAA 740
            |||||||||||
Db   721 AGCCAGCTTNGA 732

LOCUS   BM926438
DEFINITION BM926438   804 bp   mRNA   linear   EST 12-MAR-2002
            AGENCOURT_6644759 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766806
            5', mRNA sequence.
ACCESSION   BM926438
VERSION   BM926438.1   GI:19376817
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 804)
  AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE   Unpublished
  JOURNAL   NIH-MGC http://mgc.nci.nih.gov/
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12824 row: a column: 15
            High quality sequence stop: 746.

FEATURES   Location/Qualifiers
  source   1..804
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5766806"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC_122"
                  /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                  Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                  anonymous pool of 24 week female lung, 16 week female
                  spleen, and 20-22 week male spleens. Library is oligo-dT
                  primed and directionally cloned (EcoRV site is destroyed
                  upon cloning). Average insert size 1.4 kb, insert size
                  range 1-3 kb. Library is normalized and enriched for
                  full-length clones and was constructed by C. Gruber
                  (Invitrogen). Research Genetics tracking code 026. Note:
                  this is a NIH MGC Library."

BASE COUNT   165 a   238 c   243 g   125 t   33 others
ORIGIN

Query Match   98.2%;   Score 727;   DB 12;   Length 804;
Best Local Similarity   99.2%;   Pred. No. 2.9e-153;
Matches 730; Conservative   0; Mismatches   6;   Indels   0;   Gaps   0;

Qy   5 TCCGGCTGCAGCGGGGGTGAGCGGCGGCGGCAGCGGCGGCGGGGGGGGGATCCTGAGCCATGGCGGGGCGC 64
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7 TCCCGGGATAGCGGGGGTGAGCGGCGGCGGCAGCGGCGGCGGCGGGGGGATCCTGAGCCATGGGGCGGC 66

Qy   65 GCGCGACGCCATCCTGGATGCGCGTGGAGAACCTGACCGCCGGGGGGGAGCGCCAAGAAGTTCAA 124
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GCGCGCACGCCCATCCTGGATGCGCGTGGAGAACCTGACCGCCGGGGGGGAGCGCCAAGAAGTTCAA 126

Qy   125 GCTGAAGCTGCTGTCGGGTGCCGTGCGCGGAGGGCTACGGCGGCGCATCCCCCCGGGGGGCGCT 184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 GCTGAAGCTGCTGTCGGGTGCCGTGCGCGGAGGGCTACGGCGGCGCATCCCCCGGGGGGCGCT 186

Qy   185 GCTGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTA 244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 GCTGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTA 246

Qy   245 CGGCGCGGAGCTCACCGCTAAGGTGCTGCGCGGACATGGGGCTGCAGGAGGATGGGCCGGGGGCT 304
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 CGGCGCCCGGAGCTCACCGCTAAGGTGCTGCGCGGACATGGGGCTGCAGGAGGATGGGCCGGGGCA 306

Qy   305 GCTGCGCAGGCGGGCCACGCACCAGGGCTCTGGAGCGCCGCCAGCTGGGGATCCAGGCGCCCTCC 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 CGGCGCCCGGAGCTCACCGCTAACGTGCTGCGGGCCTCAGAGAGATGGGCCGGGGCA 364

```
Db   307 GCTGCAGGGCGGCCACCGCACCAGGGCTCTGGAGCCCGCGCCAGCTGGGATCCAGGCCCCTCC 366

Qy   365 TCAGTCGGCAGCCAAGCCAGGCCCTGSCACTTTATAGAACCAGCCAGCCTGCGCTTATCGC 424
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 TCAGTCGGCAGCCAAGCCAGGCCCGCACTTTATAGACCAGCCAGCCTGCGCTTATCGC 426

Qy   425 GAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTTGACGGATGA 484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTTGACGGATGA 486

Qy   485 GCAGTACCAGGCAGTGCGGGCCGGGCCCACCAACCCAGCAAGAATGCGGAAAGTTCTTCAG 544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GCAGTACCAGGCAGTGCGGGCCGGGCCCACCAACCCAGCAAGAATGCGGAAAGTTCTTCAG 546

Qy   545 TTTCACACAGCCTGGAACTGGACCTGCGAAGGACTTGCTCCTCCAGGGCCCTAAGGGAGTC 604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 TTTCACACAGCCTGGAACTGGACCTGCGAAGGACTTGCTCCTCCAGGGCCCTAAGGGAGTC 606

Qy   605 CCAGTCCTACCTGGTGGGAGGACCTCGAGCGGGCTCCTTCCCAGCAACACTCCG 664
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 CCAGTCCTACCTGGTGGGAGGACCTCGAGCGGGCTCCTTCCCAGCAACACTCCG 666

Qy   665 GTCAGCCCCTGGCAATCCACCACCAAATCATCCTGAATCTGATCTTTTTTATACACAATATAC 724
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 GTCAGCCCCTGGCAATCCACCACCAAATCATCCTGAATCTGATCTTTTTTATACACAATATAC 726

Qy   725 GAAAAGCCAGCTTGAA 740
             ||||||||||||||||
Db   727 GAAAAGCCAGCTTGNA 742
```

RESULT 3
BM910918

LOCUS   BM910918   767 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION AGENCOURT_6615922 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454497
            5', mRNA sequence.
ACCESSION   BM910918
VERSION   BM910918.1 GI:19361297
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 767)
  AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
  TITLE   Unpublished
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            NIH-MGC http://mgc.nci.nih.gov/
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1949 row: 1 column: 18
            High quality sequence stop: 641.
FEATURES   Location/Qualifiers
     source   1..767
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5454497"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit

BASE COUNT   163 a   241 c   237 g   124 t   2 others
ORIGIN
            Note: this is a NIH_MGC Library."
            (Stratagene) and Superscript II RT (Life Technologies).

Query Match   97.6%;   Score 722;   DB 12;   Length 767;
Best Local Similarity 99.7%;   Pred. No. 3.8e-152;
Matches 733;   Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy     7 CGGGCTGCAGGCGGGGTGAGGCGGCCAGGGGCGGCGGGAGCGGATCCTTGGAGCCAATGGCGCGC 66
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    17 CGGGCTGCAGGCGGGGTGAGGCGGCCAGGGGCGGCGGGAGCGGATCCTTGGAGCCAATGGCGCGC 76

Qy    67 GCGGACGGCCATCCTGGATGGGCGTGGAGAACCTGACCGGCCGGGGGGGGGGGGGGGGGCGGC 126
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    77 GCGGACGGCCATCCTGGATGGGCGTGGAGAACCTGACCGGCCGGGGGGGGGGGGGGGGGCGGC 136

Qy   127 TGAAGCTGCTGTCGGTGCCCGCTGCGCGGAGGGCGGCTACGGGCGGCGGCGGGGGGGGCTGC 186
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 TGAAGCTGCTGTCGGTGCCCGCTGCGCGGAGGGCGGCTACGGGCGGCGGCGGGGGGGGCTGC 196

Qy   187 TGTCCATGGACGGCCTTGGACCTTCACCGACAAGCTGGTCAGCTTCTACCTTGGAGACCTACG 246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 TGTCCATGGACGGCCTTGGACCTTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 256

Qy   247 GCGCCGAGCTCACCGTGCTGCGCGGCGACATGGGCCTGCGCGAGGAGGATGGCCGGCAGC 306
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 GCGCCGAGCTCACCGTGCTGCGCGGCGACATGGGCCTGCGCGAGGAGGATGGCCGGCAGC 316

Qy   307 TGCAGGCGGCCACCGACCCAGGGCTCTGGAGCCCGCGCCAGCTGGGATCCAGGCCCCTC 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 TGCAGGCGGCCACCGACCCAGGGCTCTGGAGCCCGCGCCAGCTGGGATCCAGGCCCCTC 376

Qy   367 AGTCGGCAGCCAAGCCAGGCCCTGCCACTTTATAGAACCAGCCAGCCTGCGCTTATCGCGA 426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 AGTCGGCAGCCAAGCCAGGCCCGCCACTTTATAGACCAGCCAGCCTGCGCTTATCGCGA 436

Qy   427 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCCTGACGGATGAGAGC 486
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCCTGACGGATGAGAGC 496

Qy   487 AGTACCAGGCAGTGCGGGCCGGGCCCACCAACCCAGCAAGATGCGGAAAGTTCTTCAGTT 546
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 AGTACCAGGCAGTGCGGGCCGGGCCCACCAACCCAGCAAGATGCGGAAAGTTCTTCAGTT 556

Qy   547 TCACACAGCCTGGAACTGGACCTGCGAAGGACTTGCTCCTCCAGGGCCCTAAGGGAGTCCC 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 TCACACAGCCTGGAACTGGACCTGCGAAGGACTTGCTCCTCCAGGGCCCTAAGGGAGTCCC 616

Qy   607 AGTCCTACCTGGTGGGAGGACCTGGAGCGGGCTCCTTCCCAGCAACAACTCCGGT 666
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 AGTCCTACCTGGTGGGAGGACCTGGAGCGGGCTCCTTCCCAGCAACACTCCGGT 676

Qy   667 CAGCCCCTGGCAATCCCACCACAAATCATCCT-GAATCTGATCTTTTTTATACACAAATATACG 725
            |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   677 CAGCCCCTGGCAATCCACCACCAAATCATCCTTNGAATCTGATCTTTTTTATACACAATATACG 736

Qy   726 AAAAGCCAGCTTGAA 740
            |||||||||||||||
Db   737 AAAAGCCAGCTTGAA 751
```

RESULT 4
BM549665

LOCUS   BM549665   792 bp   mRNA   linear   EST 20-FEB-2002
DEFINITION AGENCOURT_6544142 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745741
            5', mRNA sequence.
ACCESSION   BM549665
VERSION   BM549665.1 GI:18785261
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12769 row: c column: 22
            High quality sequence stop: 570.
FEATURES             Location/Qualifiers
     source          1..792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="leukocyte"
                     /clone="IMAGE:5745741"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_118"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated adult donors. Library is oligo-dT primed
                     and directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb. Library is normalized and enriched for
                     full-length clones and was constructed for
                     (Invitrogen). Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
BASE COUNT   182 a   241 c   242 g   122 t      5 others
ORIGIN
```

```
Query Match            97.5%;  Score 721.2;  DB 12;  Length 792;
  Best Local Similarity   98.8%;
  Matches  726;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
                                   Pred. No. 5.7e-152;

Qy      6 CCGGCTGCAGCCGGGGTGAGCGGGCGCACGGCGCCGGGGAATCCTGGAGCATCGGGGCGG 65
          ||  | |||| ||||||||| ||||||||| || ||||||||||||||||| |||||||||
Db     19 CCTGGATCAGCGGGGGTGACGGGCGCAGCGGCGCCGGGGAATCCTGGAGCCATCGGGGCGG 78

Qy     66 CGCGGACCGGCCATCCTTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAG 125
          |||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     79 CGCGGACCCCATCCTCTGGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAG 138

Qy    126 CTGAAGCTGCTGTCGGTGGGTGCCGCTGCGCGAGGGCGCATCCCCGCGGGGCGGCGCTG 185
          |||||||||||||||||||  |||||| ||||| |||||||||||||||||||||||||
Db    139 CTGAAGCTGCTGTCGGTGCCCGCCGCTGGGCGCATCCCCGCGGGGCGGCGCTG 198

Qy    186 CTGTCCATGGATGCGCCTTGGACTTCACCGACACAAGCTGGTCAGCTTCTACCTGGAGACCTAC 245
          |||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||
Db    199 CTGTCCATGGACGGCCTTGGACTTCACCGACACAAGCTGGTCAGCTTCTACCTGGAGACCTAC 258

Qy    246 GGGGCCCGAGCTCACCGGCTAACGTGCTGCGGCGACATGGCGCCGAGGAGGATGGCCGGCGCCAG 305
          ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db    259 GGGCGCCCAGCTCACCGGCTAACGTGGTGCCGCTGCGAGGAGTGGCCGTGGCCAG 318

Qy    306 CTTGCAGGGGGCCCACGGCACCAGGGGCTCTGGAGCCGCGCCAGCTGGGATCCAGGCCCTCCT 365
          ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CTGCAGGGGGCCCACGGCACCAGGGGCTCTGGAGCCGGATCCAGGCCCTCCT 378

Qy    366 CAGTCTCGGCGCAGCCAAGCCAGCCTGCCACTTTATAGACAGCACCACCGGGTCCGGCTTATCCCG 425
          |||| ||| ||||||||| |||||||||||||||||||||| |||| |||||||||||||||||||
Db    379 CAGTCGGCGGCAAGCCAAGCCAGCCTGCCACTTTATAGACACCACCGGGTCCTGCTTATCCCG 438

Qy    426 AGGGTCACAAACGTTGAGTGGGCTGCTGGAGGAAGGTCCTGACGGAGGATGAG 485
          ||||||||||| |||||||||||||||||||||||||||| ||||||||||||
Db    439 AGGGTCACAAACGTTGAGTGGGCTGCTGGAGGAAGGTCTGACGGAGGATGAG 498
```

```
RESULT 5
BG764161
LOCUS       BG764161          750 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602737190F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862182 5',
            mRNA sequence.
ACCESSION   BG764161
VERSION     BG764161.1  GI:14074814
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1722 row: d column: 23
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4862182"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_49"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT   165 a   232 c   235 g   118 t
ORIGIN
```

```
Query Match            96.4%;  Score 713.4;  DB 12;  Length 750;
  Best Local Similarity   99.7%;
  Matches  725;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;
                                   Pred. No. 3.2e-150;
```

```
Qy    486 CAGTACCAGGCAGTGCGGCGCCGAGCCCACCACCAAGCCCAAGACCAAGAAGAGATGCGGAAGGAGAAGCTCTTCAGT 545
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 CAGTACCAGGCAGTGCGGCGCCGAGCCCACCACCAAGCCCAAGACCAAGAAGAGATGCGGAAGGAGAAGCTCTTCAGT 558

Qy    546 TTCACACCAGCTGAACTGCAACCTGCAAGGACCTTGCTCCTCCCAGCGCCTCAAGGCGAGTCC 605
          || |||||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||
Db    559 TTCACACCAGCCCTGGAACTGGAACCTGCAAGGACCTTGCTCCTCCCAGCGCCTCAAGGGAGTCC 618

Qy    606 CAGTCCTTACCTGTTGGGAGGACGTGGGAGGCGTGAGGCTGAGGTCTCCCTTCCCAGCCCGCAACCACTCCGG 665
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    619 CAGTCCTTACCTGTTGGGAGGACCTGGGAGGCGTGAGGCTGAGGTCTCCCTTCCCAGCCCGCAACCACTCCGG 678

Qy    666 TCAGCCCCTGGGCAATCCCACCAAATCATCCTGAATCTGAATCTTTTTATACACAATATACG 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    679 TCAGCCCCTGGGCAATCCCACCAAATCATCCTGAATCTGAATCTTTTTATACACAATATACG 738

Qy    726 AAAAGCCCAGCTTGAA 740
          ||||||||  || |||
Db    739 AAAAGCCCGAAAAAAA 753
```

```
Qy    15 GCGGGGGGTGAGCGCGCGCAGCGGGGGGGGATCCTTGGAGCTGAGCTGAGCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2 GCGGGGGTGAGCGCGCGCAGCGGGGGATCCTTGGAGCTGAGCTGAGCTGAGCGCC 61

Qy    75 ATCCTGGATGCGCTGGAGAACCTGACCCGCGACCGCGGGGAGAGCTCAAGAAGTTCAAAGCTGAAGCTG 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62 ATCCTGGATGCGCTGGAGAACCTGACCCGCGACCGCGGGGAGAGCTCAAGAAGTTCAAAGCTGAAGCTG 121

Qy   135 CTGTGCGGTGCCGCTGCGGGAGGGCTACGGGGCGCCGGCGGGGCGGGGGGCGGGGGTGCTGTGTCCATG 194
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 CTGTGCGGTGCCGCTGCGGGAGGGCTACGGGGCGCCGGCGGGGCGGGGGGCGGGGGTGCTGTGTCCATG 181

Qy   195 GACGCCCTTGGACTTCACCGACAAGCTGGTTCAGCTTCTTCACCTGGAGGACCTTACGGGGCCCGAG 254
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 GACGCCCTTGGACTTCACCGACAAGCTGGTTCAGCTTCTTCACCTGGAGGACCTTACGGGGCCGAG 241

Qy   255 CTCACCCGCTAACGTGGCGTGGGCGAGGATGGGCTGCAGAGGATGTGGCCGAGGCTGAGCTTCCAGGCG 314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 CTCACCCGCTAACGTGGCGTGGGCGAGGATGGGCTGCAGAGGATGTGGCCGAGGCTGAGCTTCCAGGCG 301

Qy   315 GCCACGCACCAGGCGCTCGGGAGGCGCGGGGATCCAGGCGCCCTCCTCCTTCGGCA 374
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 GCCACGCACCAGGCGCTCGGGAGGCGCGGGGATCCAGGCGCCCTCCTCCTTCGGCA 361

Qy   375 GCCAAGCCAGGGCGGCTTATAGACCAGCAGCACCGGCGCTTATCCGGAGGGTCACA 434
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362 GCCAAGCCAGGGCGGCTGCACTTATAGACCAGCAGCACCGGCGCTTATCCGGAGGGTCACA 421

Qy   435 AACGTTGAGTGGGCGTGCGTGTACGGGAAGGTCCTGACGCGGATGAGCAGTACCAG 494
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422 AACGTTGAGTGGGCGTGCGTGTACGGGAAGGTCCTGACGCGGATGAGCAGTACCAG 481

Qy   495 GCAGTGCGGGCCCGAGCCCACCACCAGCCAAGACAAGATGCGGAAGCTCTTCAGTTTCACACCA 554
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482 GCAGTGCGGGCCCGAGCCCACCACCAGCCAAGATGCGGAAGCTCTTCAGTTTCACACCA 541

Qy   555 GCCTTGGGAACTGGACCTGCAAGGACTTGCTCCTCCTTCCAGGGCCCTAGGGAGTCCCAGTTCCTAC 614
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542 GCCTTGGGAACTGGACCTGCAAGGACTTGCTCCTCCTTCCAGGGCCCTAGGGAGTCCCAGTTCCTAC 601

Qy   615 CTGGTTGGGAGGACTTGGAGCGGAGGCTGAGGCTCCTTCCTTCCCACACACTCCGGTCAGCCCCT 674
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602 CTGGTTGGGAGGACTTGGAGCGGAGGCTGAGGCTCCTTCCTTCCCAGCACACTCCGGTCAGCCCCT 661

Qy   675 GGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTATACA-CAATATACAGAAAAGCCA 733
         |||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db   662 GGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTATACAGCAATATACAGAAAAGCCA 721

Qy   734 GCTTTGAA 740
         |||||||
Db   722 GCTTTGGA 728
```

```
Query Match              95.2%;    Score 704.8;   DB 14;   Length 725;
Best Local Similarity    99.7%;    Pred. No. 2.7e-148;
Matches     706;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy    33 GCGGGCGGGGATCCTGGAGCCATGGGGCGGCGGCGGCGGCGGCACGGCACCAGGCTCT 92
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   724 GCGGGCGGGGATCCTGGAGCCATGGGGCGGCGGCGGCACGGCACCAGGCTCT 665

Qy    93 AAACCTGACCCGCGACCGCGGGGAGAGCTCAAGAAGTTCAAAGCTGAAGCTGCGC 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   664 AAACCTGACCCGCGACCGCGGGGAGCTCAAGAAGTTCAAAGCTGAAGCTGCGC 605

Qy   153 GAGGGCTACGGGGCGCCGGCCATCCCGCGGGGCGGGCGGCTGCTGTGTCCATGGACGCCCTTGGACTTCACC 212
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604 GAGGGCTACGGGGCGCCGGCCATCCCGCGGGGCGGGCGGCTGCTGTGTCCATGGACGCCCTTGGACTTCACC 545

Qy   213 GACAAGCTGGTTCAGCTTCTTCACCTGGAGGACCTTACGGGCGCCGAGCTCACCCGCTAACGTTGGCA 272
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 GACAAGCTGGTTCAGCTTCTTCACCTGGAGGACCTTACGGGCGCCGAGCTCACCCGCTAACGTTGGCA 485

Qy   273 CGCCGACCATCGGGCCCTCGCAGGAGGATGGCCGGCGCAGCTGCAGCAGCACCCAGGGCTCT 332
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 CGCCGACCATCGGGCCCTCGCAGGAGGATGGCCGGCGCAGCTGCAGCAGCACCCAGGGCTCT 425

Qy   333 GGAGCCGGCGGCCAGCTGGATCCAGGCCTTCCTCAGTCGGCCAAGCCAAGGCCTGCCAC 392
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 GGAGCCGGCGGCCAGCTGGATCCAGGCCTTCCTCAGTCGGCCAAAACGTTGAGTGGGCGCAC 365

Qy   393 TTTATAGAACCAGCACCGGGGCTGCGCTTATCCGGAGGGTCACAAACGTTGAGTGGGCGTG 452
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TTTATAGACCAGCACCGGGGCTGCGCTTATCGGAGGGTCACAAAACGTTGAGTGGGCGGGCCAGCCC 305

Qy   453 GATGCGTCGTACGGGAAGGTCCTGACGGATGAGCGAGTACCAGGGCAGTGCGGGGGCCCAGCCC 512
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    304  GATGCTCTCTGTACGGAGAAGGTCCTGACGGAGATGAGAGTACCAGGACAGTGCGGGGCCCGGAGCCC  CGGAGCCC
```

```
Qy    513  ACCAAACCCAAGCAAGATGGCGGAAGCTCTTCAGTTTCACCAGCCTGGGAACTGACCTGC   572
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244  ACCAACCCAAGCAAGATGGCGGAAGCTTTTCAGTTTCACACCAGCCTGGGAACTGGACCTGC  185
```

```
Qy    573  AAGGACTTGCTTCCTCCAGGCCCTAAGGGAGAGTCCCAGTCTCTTACCTGGTGGAGGAACCTGGAG  632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184  AAGGACTTGCTTCCTCCAGGCCCTAAGGGAGAGTCCCAGTCTCTTACCTGGTGGAGGAACCTGGAG  125
```

```
Qy    633  CGGAGGCTGAGGCTTCCTTCCCAGCACACTCCGGTCAGCCCCTGCCCATCCCACCCAAATCA  692
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124  CGGAGGCTGAGGCTTCCTTCCCAGCACACTCCGGTCAGCCCCTGCCCATCCCACCCAAATCA  65
```

```
Qy    693  TCCTGAAATCTGATCTTTTTTATACACAATATACGAAAAAGCCAGCTTGAA  740
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64   TCCTGAAATCTGATCTTTTTTATACACAATATACGAAAAAGCCAGCTTGAA  17
```

RESULT 7

```
BG767422
LOCUS       BG767422       856 bp    mRNA    linear    EST 15-MAY-2001
DEFINITION  602741219F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4871020 5',
            mRNA sequence.
ACCESSION   BG767422
VERSION     BG767422.1  GI:14078075
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 856)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1745  row: e  column: 05
            High quality sequence stop: 696.
FEATURES             Location/Qualifiers
     source          1..856
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4871020"
                     /tissue_type="melanotic melanoma, high MDR (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_49"
                     /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT   194 a   234 c   279 g   149 t
ORIGIN
```

```
Qy    2    GCGGGGGTGAGCGGCCAGCGGGCCGGGGATCCTGAGCCATGGGGGCGGGGCGGGGGCCGGACGCC  61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    15   GCGGGGGTGAGCGGCCAGCGGGCCGGGGATCCTGGAGCCCATGGGGCGGGGGCCGGGCCGGACGCC  74
```

```
Query Match          93.4%;
Best Local Similarity  99.6%;  Score 690.8; DB 12;  Length 856;
Matches 703;  Conservative 0;  Pred. No. 4e-145;
Mismatches 2;  Indels 1;  Gaps 1;
```

```
Db    304  GCGGGGGTGAGCGGCCAGCGGGCCGGGGATCCTGAGCCATGGGGGCGGGGCGGGGGCCGGACGCC  61
```

```
Qy    75   ATCCTGGATCGGCTGGAGAGACCTGACCGCGCGGCGGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG  134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2    ATCCTGGATCGGCTGGAGAGACCTGACCGCGCGGCGGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG  61
```

```
Qy    135  CTGTCCGGTGCCGGCTGCGGCGAGGGCTACGGGGCCATCCGCGCGGGGCGCGTCGCTTGCGTCCATG  194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62   ATCCTGGATCGGCTGGAGAACCTGACCGCGCGGCGGAGGAGGCTCAAGAAGTTCAAGCTG  121
```

```
Qy    195  GACGCGCTTGGACCTCACCGACACAAGCTGGGTCAGCTTCTTACCTTGGAGGACCTTACGGAG  254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122  CTGTCGGTGCCGGCTGCGGCGAGGGCTACGGGGCCATCCGCGCGGGGCGCTGCGTCCATG  181
```

```
Qy    255  CTCACCCGCTAACGTGCTGCCCGACCATGGGCCTGCAGGAGATGGGGCGGGGGCGGGG  314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182  GACGCGCTTGGACCTCACCGACACAAGCTGGGTCAGCTTCTTACCTGGGGAGCCTTA  241
```

```
Qy    315  GCCACCGCACCAGGGGTTCTGGAGCCCGCGCCCGCGGGGGTCAGGCCCCTCCTCCTTCAGTCGGCA  374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242  CTCACCCGCTAACGTGCTGCCCGACCATGGGCCTGCAGGAGATGGGGCGGGGGCGGGG  301
```

```
Qy    375  GCCAAGCCAGGCCCTGCACTTTATAGATACAGCCACCACGCCCGGCGCTTATCGCGAGGGGTCACA  434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302  GCCACCGCACCAGGGGTCTGGAGCCCGCGCCCTCTTCAGTCGGCA  361
```

```
Qy    435  AACGTTGAGTGGGCTCCTTGGAGTGGCTTCTTGTACGGGGAAGGTCCTGACGGATGGAGCAGTACCAG  494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362  GCCAAGCCAGGCCCTGCACTTTATAGATACAGCCACCACGCCCGGCGCTTATCGCGAGGGGTCACA  421
```

```
Qy    495  GCCAGTGCGGCCGGAGCCCACCACGCCCAAGCAAGATGCCGGAAGGTCCTGACGGATGGAGCAGTACCAG  554
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422  AACGTTGAGTGGCTGCTTGGATGCTCCGTGTACGGGGAAGGTCCTGACGGATGGAGCAGTACCAG  481
```

```
Qy    555  GCCCTGGAACTGGACCTGCAAGGAGACTTGCTCCTCCCAGGCCCTAAGGGAGAGTCCCAGTCCTTAC  614
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482  GCAGTGCGGCCGGCGCCACCCAACCCCAAGCAAGATGCCGGAAGGTCCTTCAGTTTCACACA  541
```

```
Qy    615  CTGGTGGAGGACCTGGAGCCGGAGCTGAGGGCTCCTTCCTCTCCCAGGCCCTAAGGGAGCTCTTCAGTTTCACGCCCT  674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542  GCCCTGGAACTGGACCTGCAAGGAGACTTGCTCCTCCCAGGCCCTAAGGGAGAGTCCCAGTCCTAC  601
```

```
Qy    675  GGC-AATCCCACCCAAATCATCCCTGAATCTGAATCTGATCTTTTTTATACACAA  719
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602  CTGGTGGAGGACCTGGAGCCGGAGCTCCTTCCCTCCCAGCGACACTCCGGTCAGCCCCT  661
```

```
Db    662  GGCTAATCCCACCCAAATCATCCCTGAATCTGAATCTGATCTTTTGTATACACAA  707
```

RESULT 8

```
BG337806
LOCUS       BG337806       748 bp    mRNA    linear    EST 27-FEB-2001
DEFINITION  6024357009F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4553378 5',
            mRNA sequence.
ACCESSION   BG337806
VERSION     BG337806.1  GI:13144342
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 748)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1251  row: b  column: 03
```

FEATURES
```
     source          1..748
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4553378"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC 46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GCCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT      153 a    239 c    238 g    118 t
ORIGIN
```

Query Match   93.1%; Score 688.6; DB 10; Length 748;
Best Local Similarity   98.5%; Pred. No. 1.2e-144;
Matches 716; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
Qy       15  GCCGGGGTGAGCGCGGCGCACGGCGCCGGGGGATCCTGAGCCATGGGCCGCGGCGACGCC   74
             | ||||||||||||| | ||| | |||  || ||||||| |||||||||||||||||||||
Db        2  GCGGGGGTGAGCGGCGGCAGCCGGCGGATCCTGGAGCCATGGGCCGCGGCGACGCC   61

Qy       75  ATCCTGGATGCGCTGGAGAACCTGACCGCGCCGGGAGGAGCTCAAGAAGTTCAAGCTGAAGCTG  134
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       62  ATCCTGGATGCGCTGGAGAACCTGACCGCGCCGGGAGGAGCTCAAGAAGTTCAAGCTG  121

Qy      135  CTGTCGGGTGCCGCTGCGCGGAGGGCTTACGGGCGCCATCCCGCGGGGCGGCGGCTGTCCATG  194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      122  CTGTCGGGTGCCGCTGCGCGGAGGGCTTACGGGCGCCATCCCGCGGCTGCGTCCATG  181

Qy      195  GACGCCCTTGGACCTCACCCACAACAACCTGGTCAGCTTCTTCACCTGGAGACCTTACGGGCGGGCCGAG  254
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      182  GACGCCCTTGGACCTCACCGACAACAAGCTGGTCAGCTTCTTCACCTGGAGACCTTACGGGCCCCGAG  241

Qy      255  CTCACCGCTAACGTGCTGGGCGGCGACATGGGCCTGCAGGAGATGGCCGGCGAGCGCTGCAGGCG  314
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      242  CTCACCGCTAACGTGCTGGGCGACATGGGCCTGCAGGAGATGGCCGGCGAGCGCTGCAGGCG  301

Qy      315  GCCACGCACCAGGGGCTCTGGAGCCCGCGCCAGCTGGGGATCCAGGGCCCCTCCTCAGTCGGCA  374
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      302  GACACGCACCAGGGCTCTGGAGCGCGCCAGCTGGGGATCCAGGGCCCCTCCTCAGTCGGCA  361

Qy      375  GCCAAGCCAGGCCTGCACTTTATAGACACAGCACCACCGGGGCTGCGCGCTTATCCGGAGGGGTCACA  434
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      362  GCCAAGCCCAGGCCTGCACTTTATAGACACAGCACCACCGGGCTTATCGGCAGCAGTCACA  421

Qy      435  AACGTTGAGTGGGCTGGATGCCTCGTGTACGGGGAAGGTCCTGACGGATGAGCAGTACCAG  494
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      422  AACGTTGAGTGGGCTGGATGCCTCGTGTACGGGGAAGGTCCTGACGGATGAGCAGTACCAG  481

Qy      495  GCAGTGCGGCGCCGAGCCCACCACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACCA  554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      482  GCAGTGCGGCGCCCACCACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACCA  541

Qy      555  GCCTGGAACTGGACCTTGCTCAAGGACTTGCTCCTTCCCAGGGAGTCCCACTCCTAC  614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      542  GCCTGGAACTGGACCTTGCTCAAGGACTTGCTCCTTCCCAGGGAGTCCCACTCCTAC  601

Qy      615  CTGGTGGGAGGAACTGGACCTGCAAGGGACTTGCTCCTTCCCAGCAACACTCCGGTCAGCCCCT  673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      602  CTGGTGGGAGGAACTGGACCTGGAGTTGGAGGGAGTGCCAGCAACACTCCGGTCAGCCGCCCCT  661

Qy      674  TGGCAATCCCACCAAAATCATCCTGGAATCTGATCTTTTATACACAATATACGGAAAAGCCA  733
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      662  TGTCAATCCCAGCAGAGATCATCCTGGAATCTGATC-TTATACACACAATATACGGAGAAGCCA  720

Qy      734  GCTTGAA  740
             ||||| |
Db      721  GCTTGCA  727
```

RESULT 9
AI148558/c

```
LOCUS       AI148558                678 bp    mRNA    linear   EST 28-OCT-1998
DEFINITION  qc68g10.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1714818 3' similar to TR:O15553 O15553 PYRIN. i, mRNA
            sequence.
ACCESSION   AI148558
VERSION     AI148558.1  GI:3677027
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbc-r@mail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 804   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 402.
FEATURES
     source          1..678
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1714818"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGACGGGGCGATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT      114 a    202 c    217 g    145 t
ORIGIN
```

Query Match   90.6%; Score 670.4; DB 9; Length 678;
Best Local Similarity   99.9%; Pred. No. 1.5e-140;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       69  GACGGCCATCCTGGATGCGCTGGAGAACCTGACCGCGCCGGGAGGAGCTCAAGAAGTTCAAGCTG  128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      678  GACGGCCATCCTGGATGCGCTGGAGAACCTGACCGCGCCGGGAGGAGCTCAAGAAGTTCAAGCTG  619

Qy      129  AAGCTGCTGTCGGGTGCCGCTGCGCGGAGGGCTTACGGGCGCCATCCCGCGGGGCGGCGGCTG  188
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      618  AAGCTGCTGTCGGGTGCCGCTGCGCGGAGGGCTTACGGGCGCCATCCCGCGGGGCGGCGGCTG  559

Qy      189  TCCATGGACGCCCTTGGACCTCACCGACAACAACCTGGTCAGCTTCTTCACCTGGAGACCTTACGGC  248
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      558  TCCATGGACGCCCTTGGACCTCACCGACAACAAGCTGGTCAGCTTCTTCACCTGGAGACCTTACGGC  499

Qy      249  GCCGAGCTCACCGCTAACGTGCTGGGCGACATGGGCCTGCAGGAGATGGCCGGCGAGCGCTG  308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      498  GCCGAGCTCACCGCTAACGTGCTGGGCGACATGGGCCTGCAGGAGATGGCCGGCGAGCAGCTG  439
```

```
Qy    309 CAGGCCGGCACGCACCAGGGACTCTTGGAGCGCCGCGCGCCGCAGCTTGGGGATCCAGCCCCTTCCCTCAG 368
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 CAGGCCGGCACGCACCAGGGGCTCTTGGAGCGCCGCGCGCCGCAGCTTGGGGATCCAGCCCCTTCCCTCAG 379

Qy    369 TCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACACCCGGGCGCTGCGCTTATCGGCGAGG 428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 TCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACACCCGGGCGCTGCGCTTATCGGCGAGG 319

Qy    429 GTCACAAACGTTGAGTGGCGTCGGATGCTCTGTACGGGATGAGCAG 488
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    318 GTCACAAACGTTGAGTGGCTGCTCGTACGGGGAGAAGGGTCCTGACGGGATGAGCAG 379

Qy    489 TACCAGGCAGTGCGGGCCCAGCCCACCAACCAAGCAAGATGCGGGAAGCTCTTCAGTTTC 548
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 TACCAGGCAGTGCGGGCCCAGCCCACCAACCAAGCAAGATGCGGGAAGCTCTTCAGTTTC 199

Qy    549 ACACCAGCCTGGAACTGGACCTGCAAGGACTGCTGCTCCTCCCAGGCCCTAAGGGAGTCCCAG 608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 ACACCAGCCTGGAACTGGACCTGCAAGGACTGCTGCTCCTCCCAGGCCCTAAGGGAGTCCCAG 139

Qy    609 TCCTACCCTGGTGGAGGACGGGACGGCTCTTCTTCCCAGCAACACTCCGGTCA 668
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    138 TCCTACCCTGGTGGAGGACGGGACGGCTCTTCTTCCCAGCAACACTCCGGTCA 79

Qy    669 GCCCCTGGGCAATCCCACCAAAATCATCCTCGAAATCGATCTTTTTTATACACAATATACGAAA 728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 GCCCCTGGGCAATCCCACCAAAATCATCCTCGAAATCGATCTTTTTTATACACAATATACGAAA 19

Qy    729 AGCCAGCTTGAA 740
          ||||||||||||
Db     18 AGCCAGCTTGAA 7
```

RESULT 10
BM051141
LOCUS   BM051141   741 bp   mRNA   linear   EST 07-NOV-2001
DEFINITION   603634063F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:5424590 5',
             mRNA sequence.
ACCESSION   BM051141
VERSION   BM051141.1   GI:16780408
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 741)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM1885 row: n column: 15
             High quality sequence stop: 643.
FEATURES   Location/Qualifiers
     source   1..741
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5424590"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC 43"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG (G). Library constructed by Ling Hong

                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT   139 a   243 c   237 g   122 t
ORIGIN

Query Match   90.2%;   Score 667.6; DB 12;   Length 741;
Best Local Similarity   97.1%;   Pred. No. 6.4e-140;
Matches 701; Conservative   0; Mismatches   19; Indels   2; Gaps   2;

```
Qy      7 CGGGCTGCAGGGGGGTGAGCGGGCGGCGGCCAGGGCGGCCGGCGCCGGGGATCCTGGACCCATGGGGCGC 66
          |||||||||||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     21 CGGGCTGCAGCGGGGGTTGAGCGGGCGGCGGCCAGGGCGGCCGGCGTCCTGGACCCATGGGGCGGGGGGC 80

Qy     67 GCGACGGCCATCCTGGATGCGCTGGAGAACCTGACCGGCCGGGGGGCTCAGGAAGTTCAAGC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GCGACGGCCATCCTGGATGCGCTGGAGAAACCTGACCGGCCGAGTAGTCAAGAAGTTCAAGC 140

Qy    127 TGAAGCTGCTGTCGGTGCCCGCTGCGGCGAGGGGCTACGGGGGCGGCGCGATCCCGGGCGTGC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 TGAAGCTGCTTGTCGGTTGCCCGCTGCGGCGAGGGGCTACGGGGGCGGCCGGCGCTGC 200

Qy    187 TGTCCATGGACGGCCTTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACTTACG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 TGTCCATGGACGGCCTTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGACCTACG 260

Qy    247 GCGGCGGAGCTTCACCCGCCTAACGTGCTGCGGCCGCGGCCAGCATGGGCCTGCAGGAGGATGGCCCAGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 GCGGCCGGAGCTTCACCCGCCTAACGTGCTGCGGCCGCGGACCAGCATGGGCCTGCAGGAGGATGGCCCAGC 320

Qy    307 TGCAGGCGGCCGGCCACCAGGGCGCTCTGGAGCGGCGGCCGCCAGCTGGGATCCAGGCCCCTCTC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 TGCAGGCGGCCGGCCACCAGGGCGCTCTGGAGCGGCGGCCGCCAGCTGGGATCCAGGCCCCTCTC 380

Qy    367 AGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACACCCGGGCGCTGCCCTTATCGGCGA 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    381 AGTCGGCGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACACCCGGGCGCTGCGCCTTATCGGCGA 440

Qy    427 GGGTCACAAACGTTGAGTGGCTGCTCGTCGTACGGGGAGGTCCTGACGGATGAGC 486
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 GGGTCACAAACGTTGAGTGGCTGCTCGTCGTACGGGGAGGTCCTGACGGATGAGC 500

Qy    487 AGTACCAGGCAGTGCGGGCCGCGGAGCCCACCAACCCAAGCAAGATGCGGAAGCTCTTCAGTT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 AGTACCAGGCAGTGCGGGCCGCGGAGCCCATCATACCAAGCAAGATGGGGAAGCTCTTCAGTT 560

Qy    547 TCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCCCAGGCCCTAAGGGAGTCCC 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 TCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCCCAGGCCCTAAGGGAGTCCC 620

Qy    607 AGTCCTTACCTGGGTGGAGGACCT-GGAGCGGAGGCTTCCTTCCCAGCAACACTCCGG 665
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    621 AGTCCTTACCTGGGTGGAGGACCTGGGAGGCGGAGGCTTCCTTCCCAGCAACACTCCG 680

Qy    666 GCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACG 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    681 TCCAGCCCTTGGCAAT-CCACCAAAATCATCCTGAATCTTCATACCCCCACTCCG 739

Qy    726 AA 727
          ||
Db    740 AA 741
```

RESULT 11
BM972785/c
LOCUS   BM972785   676 bp   mRNA   linear   EST 20-FEB-2003
DEFINITION   UI-CF-EC1-abr-k-24-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
             UI-CF-EC1-abr-k-24-0-UI 3', mRNA sequence.
ACCESSION   BM972785
VERSION   BM972785.1   GI:19590376
KEYWORDS   EST.
SOURCE   Homo sapiens (human)

```
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 676)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
             discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      9704477
PUBMED       8889548
COMMENT      Contact: McCray, PB
             McCray Lab
             University of Iowa
             2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
             Tel: 319 356 4866
             Fax: 319 356 7171
             Email: paul-mccray@uiowa.edu
             Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
             Genetics (www.resgen.com) or from Open Biosystems
             (www.openbiosystems.com).
             Seq primer: M13 FORWARD
             POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..676
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EC1-abr-k-24-0-UI"
                     /tissue_type="Lung"
                     /dev_stage="Adult and Fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EC1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EC1 is a normalized cDNA library containing the
                     following tissue(s): Normal lung from adult and from fetal
                     day 64, day 87, week 19 and week 42. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806,1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AAGTGCTTAC.
                     TAG LIB=UI-CF-EC1
                     TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
                     and 380-383
                     TAG_SEQ=AAGTGCTTAC"
BASE COUNT      112 a    199 c    211 g    154 t
ORIGIN
Query Match     89.2%;  Score 660;  DB 12;  Length 676;
Best Local Similarity   100.0%;  Pred. No. 3.2e-138;
Matches  660;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      81 GATCGCGCTGGAGAACCTGACCGCCCAGGAGACTCAAGAAAGTTCAACCTGAAGCTGCTGTCG 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     676 GATCGCGCTGGAGAACCTGACCGCCCAGGAGACTCAAGAAAGTTCAACCTGAAGCTGCTGTCG 617

Qy     141 GTGCCGGCTCGGCGGAGGGCTACGGGCGGCGATCCCGCGCGGGGGCGGGGCCGG 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     616 GTGCCGGCTCGGCGGAGGGCTACGGGCGGCGATCCCGCGCGGGGGCGGGGCCGG 557

Qy     201 TTGGACCTCACCGACAAGCTGGTCAGCTTCTTACCTGGAGACCTACGGCGCCGCGGAGCTCACC 260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     556 TTGGACCTCACCGACAAGCTGGTCAGCTTCTTACCTGGAGACCTACGGCGCCGCGGAGCTCACC 497
```

```
RESULT 12
LOCUS       BE908204         675 bp    mRNA         linear   EST 20-OCT-2000
DEFINITION  601S004S1F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902319 5',
            mRNA sequence.
ACCESSION   BE908204
VERSION     BE908204.1  GI:10402543
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 675)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9704  row: j  column: 16
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1..675
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3902319"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.1 kb. Library constructed by Life
```

```
Qy     261 GCTAACCTGCTGCGCGACATGGGCGTGGGCCTGCAGGAGGATGGCCGGGCAGGCTGCAGGGCGGGCCACG 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     496 GCTAACGGTGCTGCGCGCGCACATGGGCCTGCAGGACATGGCCGAGAGGATGGCCGGGCAGGGCGGGCCACG 437

Qy     321 CACCAGGGCTCTGGAGCCGCGGCCACTGGGATTCCAGGCGTGGGCCCCTCCTCCAGTCTCGGCGGCCAAG 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 CACCAGGGCTCTGGAGCCGCGGCCACTGGGATTCCAGGCGTGGGCCCCTCCTCCAGTCTCGGCGGCCAAG 377

Qy     381 CCAGGGCCTGCACTTTATAGACCAGCACCCGGCGCTGCCGCCTTATCGGCCGAGGGTCACAAACGTT 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 CCAGGGCCTGCACTTTATAGACCAGCACCCGGCGCTGCCGCCTTATCGGCCGAGGGTCACAAACGTT 317

Qy     441 GAGTGGCGCTGCTGGATGCTCTGTACGGGGAAGGTCCTGACGGATGAGCAGTACCAGGCAGTG 500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 GAGTGGCGCTGCTGGATGCTCTGTACGGGGAAGGTCCTGACGGATGAGCAGTACCAGGCAGTG 257

Qy     501 CGGGGCCGAGCCCCACCAAACCCAAGCAAAGATGCGGGAAGCTCTTCAGTTTCACACCAGCCTGG 560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 CGGGGCCGAGCCCCACCAAACCCAAGCAAAGATGCGGGAAGCTCTTCAGTTTCACACCAGCCTGG 197

Qy     561 AACTGGACCTGCAAGGACTTGCTCCTCCAGGGCCCTAAGGAGGAGTCCCAGTGTCCTACCTGGTG 620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     196 AACTGGACCTGCAAGGACTTGCTCCTCCAGGGCCCTAAGGAGGAGTCCCAGTGTCCTACCTGGTG 137

Qy     621 GAGGACCTGGAGCGGAGCTGAGGCTCCTTCCCAGCAACACTCCGGTTCAGCCCCTGGCAAT 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136 GAGGACCTGGAGCGGAGCTGAGGCTCCTTCCCAGCAACACTCCGGTTCAGCCCCTGGCAAT 77

Qy     681 CCCACCAAAATCATCCTTGAATCTGATCTTTTTTATACAACAATATACGAAAAAGCCAGCTTGAA 740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 CCCACCAAAATCATCCTTGAATCTGATCTTTTTTATACAACAATATACGAAAAAGCCAGCTTGAA 17
```

```
BASE COUNT      151 a      214 c      198 g      112 t
ORIGIN

                                Technologies. "

Query Match            88.4%;   Score 654.2;  DB 10;  Length 675;
Best Local Similarity  99.5%;   Pred. No. 6.9e-137;
Matches  656;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy     82 ATGCGGCTGGAGAACCTGACCGGCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGTCGG 141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGCGGCTGGAGAACCTGACCGGCCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGTCGG 60

Qy    142 TGCCGGCTGCCGCGAGGGCTACGGGCGGCCATCCCGCGGGGGCGCTGCTGGTCCATGGACGCGCCT 201
          ||||||||| ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     61 TGCCGGCTGCGGCGAGGGCTACGGGCGGCCATCCCGCGGTGTCCATGGACGCCT 120

Qy    202 TGGACTTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGGCGGCCGAGCTCACCG 261
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGACTTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGGCGGCCGAGCTCACCG 180

Qy    262 CTAACGTGCTGCCGCGACATGGGGCCTGCGAGGAGAGATGCCGGCCGCCACGC 321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTAACGTGCTGCCGCGACATGGGGCCTGCGAGGAGAGATGCCGGCCGCCACGC 240

Qy    322 ACCAGGGGCTTCTGAGCCGCCGCCAGCTGGGATCCAGGCCCCTCCTCAGTGCGGCAGCCAAGC 381
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACCAGGGCTTCTGAGCCGCCGCCAGCTGGGATCCAGGCCCCTCCTCAGTGCGGCAGCCAAGC 300

Qy    382 CAGGCCCTGCACTTTATAGACACCAGCACCGGGCTGGCCTTATCGCGAGGGTCACAAACGTTG 441
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGGCCCTGCACTTTATAGACACCAGCACCGGGCTGGCCTTATCGCGAGGGTCACAAACGTTG 360

Qy    442 AGTGGCTGCTGGATGCTCTGTACGGGGGAAGGTTCTTGACGGATGAGCAGTACCAGGCAGTGC 501
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGTGGCTGCTGGATGCTCTGTACGGGGGAAGGTTCTTGACGGATGAGCAGTACCAGGCAGTGC 420

Qy    502 GGGCCGGAGCCCACCAACCCAAGCAAAGCTCTTCAGTTTCACACCAGCCTGGA 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGCCGGAGCCCACCAACCCAAGCAAAGCTCTTCAGTTTCACACCAGCCTGGA 480

Qy    562 ACTGGACCTGCAAGGACTTGCTCCTCCCAGGGCCCTAAGGGAGTCCCAGTCCTACCTGGTGG 621
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACTGGACCTGCAAGGACTTGCTCCTCCCAGGGCCCTAAGGGAGTCCCAGTCCTACCTGGTGG 540

Qy    622 AGGACCTGGAGCCGGAGGTGAGGGCTCCTCCCACCACTCCCGGTCACCCCTTGGGCAATC 681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGACCTGGAGCCGGAGGTGAGGGCTCCTCCCACCACTCCCGGTCACCCCTTGGGCAATC 600

Qy    682 CCACCAAAATCATCCTGAATCTGATCTTTTTATACACAATAATACGAAAAAGCCAGTTGAAA 740
          ||||||||||||||||||||||||||||||||| ||||||||||||||||| ||
Db    601 CCACCAAAATCATCCTGAATCTGATCTTTTTATACACAATAATACGAAAAAGCCAGTTGAAA 659
```

```
RESULT 13
LOCUS       BG255521              744 bp    mRNA    linear    EST 13-FEB-2001
DEFINITION  602367671F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476088 5',
            mRNA sequence.
ACCESSION   BG255521
VERSION     BG255521.1  GI:12765259
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM103020 row: m column: 17
            High quality sequence stop: 661.
FEATURES             Location/Qualifiers
     source          1..744
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4476088"
                     /tissue_type="adenocarcinoma, cell line"
                     /clone_lib="NIH_MGC_91"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb; Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      159 a      227 c      238 g      120 t
ORIGIN

Query Match            87.5%;   Score 647.4;  DB 10;  Length 744;
Best Local Similarity  97.0%;   Pred. No. 2.2e-135;
Matches  713;  Conservative  0;  Mismatches  16;  Indels  6;  Gaps  5;

Qy      8 GGCTGCAGCGCGGGGTCAGCGCGGCCAGCGCGCGGGGGATCCTGGGACGCCATGGACGCGCGCG 67
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGCTGCAGCGGGGGTTGAGCGCGGCCAGCGCGCGGGGGATCCTGGGACGCCATGGACGCGCCG 60

Qy     68 CGAGGCCATCCTGATGGCCTGCGAGGAACCTGACCGCCGAGGGAGTTCAAGCT 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGAGGCCATCCTGATGGCCTGCGAGGAACCTGACCGCCGAGGGAGTTCAAGCT 120

Qy    128 GAAGCTGCTGTCGGTGCCGGCCGCTGCGCGGAGGGCTACGGGCGGCCATCCCGCGGTGCTGCT 187
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAAGCTGCTGTCGGTGCCGGCCGCTGCGCGGAGGGCTACGGGCGGCCATCCCGCGGTGCTGCT 180

Qy    188 GTCCATGGACGCGCCTTGGACCTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGG 247
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCCATGGACGCCTTGGACCTCACCGACCAAGCTGGTCAGCTTCTACCTGGAGACCTACGG 240

Qy    248 CGGCGGAGCTCACCGCTAACGTGCTGCCGCGACACATGGGGCCTGCGAGGAGAGATGGCGGCAGCT 307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGGCCGAGCTCACCGCTAACGTGCTGCCGCGACACATGGGGCCTGCGAGGAGAGATGGCGGCAGCT 300

Qy    308 GCAGGCGGGGCCCACCGGACCAGGGCTTCTGAGCCGCCGCCAGCTGGGATCCAGGCCCCTCCTCA 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCAGGCGGGGCCACCGGACCAGGGCTTCTGAGCCGCCGCCAGCTGGGATCCAGGCCCCTCCTCA 360

Qy    368 GTCGGCAGCCAAGCCAGGCCAGGCGCTGCACTTTATAGACACCAGCACCGGGGCTGCGGCCGCAG 427
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTCGGCAGCCAAGCCAGGCCAGGCTGCACTTTATAGACACCAGCACCGGGCTGCCGGCCGCAG 420

Qy    428 GGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGGGAAGGTTCCTGACGGATGAGCA 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGTCACAAACGTTGAGTGGCTGCTGGATCGTCTCGTTACGGGGGAAGGTTCCTGACGGATGAGCA 480

Qy    488 GTACCAGGCAGTGCGGGCCGGAGCCCACCAACCCAAGCAAGAGATGCGGGGAAGCTCTTCAGTTT 547
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTACCAGGCAGTGCGGGCCGGAGCCCA-CAACCCAAGCAAGATGGGGAAAGCTCTTCAGTTT 539

Qy    548 CACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCCAGGGCCCTAAGGGGAGTCCCA 607
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CACACCAGCCT-GAACTGGACCTGCAAGGACTTGCTCCTCCCAGG--CCTAAGGGGAGTCCCA 597

Qy    608 GTCCTACCTGGTGGAGGACCTGGAGCCGGAGGCTGAGGCT-CTTCCAGCAACACTCCGGGTC 667
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    598 GTCCTACCTGGTGGAGGACCTGGAGCGGAGGCT-CTTCCAGCAACACTCCGGGTC 656

Qy    668 AGCCCCTGCAATCCCACCAAAATCATCCTGAAT--CTGATCTTTTTATACACAATAATACG 725
          ||||||||||||||||||||||||||||||||||
Db    ...
```

RESULT 14

```
Db      657 AGCCCCGGGAAATCCCACCAAAAATATCCTGGATTCTGATCTTTTTTATACACAATATACG
            ||||||  ||||||||       ||         ||||||||||||||||||||||||||
Qy      726 AAAAGCCAGCTTGAA                                              740
            ||||||| ||||| ||
Db      717 AAAAGCCAGTTGAAA 731
```

LOCUS   BE909218
DEFINITION 601501478F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3903498 5',
            mRNA sequence.
ACCESSION   BE909218
VERSION   BE909218.1 GI:10404770
KEYWORDS   EST.
SOURCE   Homo sapiens
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 687)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC_Clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9707 row: k column: 19
            High quality sequence stop: 673.
FEATURES   Location/Qualifiers
     source   1..687
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3903498"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib=NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.1 kb. Library constructed by Life
                     Technologies."
BASE COUNT   140 a   219 c   214 g   114 t
ORIGIN

```
Qy       48 GGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGTCGGTCGGCCGCCTGCGGGCGCGACGCCGC  107
            ||||||||||||||||||||||||||||||| ||||| |||||||||||||| ||||||||
Db        1 GGAGCTCATGGGGCGCGC--CGCGCGACGCCGCATCCCTGCATGCGCCTGGACGACGGCCGA  60

Qy      107 GGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGTCGGTCGGCCGCGAGGGCTTACGGGGCG    166
            ||||||||||||||||||||||||||| ||||||||||||| |||||||||||||||||
Db       61 GGAGCTCAAGAAGTTCAAGCTGCTGTCGGTCGCCCTGCGCGAGGGGCTTACGGGGGC     120

Qy      167 CATCCCGCCGGCGGGCGCCGCTGCTGTCCATGGACGGCCGTCGACCAAGCTGGGTCAG     226
            ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CATCCCCGCGGCGGGCGCCGCTGCTGTCCATGGACGGCCTTGGACCAAGCTGGGTCAG     180

Qy      227 CTTCTACCTGGAGACCTACGGGCGCCGAGCTCACCCGGCTCACCCGGCACATGGGGCCT     286
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CTTCTACCTGGAGACCTACGGGCGGCGAGCTCACGCGGCTCACCCGGCACATGGGGCCT     240

Qy      287 GCAGGAGATGGGCGCGGGCGGGCCAGCGACCCAGGCGGCCACCAGGGGCTCTGGAGCCGCCAGC    346
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 15

```
Db      658 TTTTTATACACAATATACGAAAAGC 682
            |||||||||||||||||||||||||
Qy      707 TTTTTATACACAATATACGAAAAGC 731
            |||||||||||||||||||||||||
Qy      601 --TTCCCAGCACACTCCGGTCAGCCCCCTGG--AATCCCACCAAAATATCATCCCTGAAATCTGATC  657
```

LOCUS   BM998354/c
DEFINITION UI-H-DT1-awc-p-21-0-UI.s1 NCI_CGAP_DT1 Homo sapiens cDNA clone
            IMAGE:5887748 3', mRNA sequence.
ACCESSION   BM998354
VERSION   BM998354.1 GI:19723255
KEYWORDS   EST.
SOURCE   Homo sapiens
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 638)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..638
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5887748"
                     /tissue_type="Metastatic Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_DT1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DT1 is a normalized cDNA library containing the
                     following tissue(s): Metastatic Chondrosarcoma in Lung.
                     The library was constructed according to Bonaldo, Lennon
```

and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is AACTGTTCGG.
TAG_LIB=UI-H-DT1
TAG_TISSUE=lung metastatic chondrosarcoma
TAG_SEQ=AACTGTTCGG"

BASE COUNT   108 a   185 c   201 g   144 t
ORIGIN

Query Match   83.4%;   Score 617.2;   DB 12;   Length 638;
Best Local Similarity   99.5%;   Pred. No. 1.3e-128;
Matches   619;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy   119 GTTCAAGCTGAAGCTGCTCGTCGGTGCCCGCTGCGGAGGGCCTACGGGCGGCATCCCGCGGGG 178
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   638 GTCAAAGCTGAAGCTGCTCGTCGGTGCCCGCTGCGGAGGGCCTACGGGCGGCATCCCGCGGGG 579

Qy   179 CGCGCTGCTGTCCATGACGCCGTTGGGACCTCACCGACAAAGCTGGTCAGCTTCTTACCTGGA 238
             ||||||||||||||| |||| |||||||||||||||| ||||||||||||||||||||||||
Db   578 CGGGCTGCTGTCCATGGACGCCTTGGGACCTCACCGACAAGCTGGTCAGCTTCTTACCTGGA 519

Qy   239 GACCTACGGCGCCGCGAGCTCACCCGTAACGTGCTGCGCGCCACATGGGGCCTGCAGGAGATGGC 298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   518 GACCTACGGCGCCGCGAGCTCACCCGTAACGTGCTGCGCGCCACATGGGGCCTGCAGGAGATGGC 459

Qy   299 CGGGCAGCTGCAGGCGGGCCACGGGCCACGGACCAGGGCTCTGGAGCGGCGCCAGCTGGGATCCAGGC 358
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   458 CTGGCAGCTGCAGGCGGGCCACGGGCCACGGACCAGGGCTCTGGAGCGGCGCCAGCTGGGATCCAGGC 399

Qy   359 CCCTCCTCAGTCGGCAGCCCAAGCCAGGGCTCGCACTTTATAGACCACGCACCGGGCTGCGCT 418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   398 CCCTCCTCAGTCGGCAGCCCAAGCCAGGGCTCGCACTTTATAGACCACGCACCGGGCTGCGCT 339

Qy   419 TATCGCGAGGGGTCACAAACGTTGAGTGGCTCTGGGGATCGCTCTCGTACGGGAAGGTCCTGAC 478
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   338 TATCGCGAGGGGTCACAAACGTTGAGTGGCTCTGGGGATCGCTCTCGTACGGGAAGGTCCTGAC 279

Qy   479 GGATGAGGCAGTACCAGGCCAGTGCGGCCGAGCCCACCCAAACCCAAGCAAAGATGCGGAAGCT 538
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   278 GGATGAGGCAGTACCAGGCCAGTGCGGCCGAGCCCACCCAAACCCAAGCAAAGATGCGGAAGCT 219

Qy   539 CTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCTCCAGGCCCTAAG 598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   218 CTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCTCCAGGCCCTAAG 159

Qy   599 GGAGTCCCAGTCCTACCTGGGTGGAGCGGAGGCTCGAGGGCTCCTTCCCAGCAAC 658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158 GGAGTCCCAGTCCTACCTGGGTGGAGCGGAGGCTCGAGGGCTCCTTCCCAGCAAC 99

Qy   659 ACTCCGGTCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTATACACA 718
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    98 ACTCCGGTCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTATACACA 39

Qy   719 ATATACGAAAAGCCAGCTTGAA 740
             ||||||||||||||||||||||
Db    38 ATATACGAAAAGCCAGCTTGAA 17
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 03:02:31

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcggggg..........atacgaaaagccagcttgaa 740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 56978 seqs, 22069156 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

; Search time 39.6087 Seconds
; (without alignments)
; 8246.264 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match Length	DB	ID	Description
1	740	100.0	740	4	US-09-340-620A-48	Sequence 48, Appl
2	585	79.1	585	4	US-09-340-620A-50	Sequence 50, Appl
3	336	45.4	777	4	US-09-340-620A-60	Sequence 60, Appl
4	316.2	42.7	579	4	US-09-340-620A-62	Sequence 62, Appl
5	53	7.2	909	4	US-09-252-991A-2958	Sequence 2958, Ap
6	53	7.2	1218	4	US-09-252-991A-3146	Sequence 3146, Ap
7	53	7.2	1380	4	US-09-252-991A-2822	Sequence 2822, Ap
8	53	7.2	1500	4	US-09-252-991A-3050	Sequence 3050, Ap
9	53	6.8	2079	4	US-09-252-991A-3097	Sequence 3097, Ap
10	50.6	6.7	33529	4	US-09-144-085-3	Sequence 3, Appl
11	49.8	6.7	50937	3	US-09-428-517-1	Sequence 1, Appl
12	49.2	6.6	2610	4	US-09-252-991A-1107	Sequence 1107, Ap
13	49	6.6	3114	4	US-09-252-991A-977	Sequence 977, App
14	49	6.6	3195	4	US-09-252-991A-937	Sequence 937, App
15	48.8	6.6	468	4	US-09-252-991A-10961	Sequence 10961, A
16	48.8	6.6	795	4	US-09-252-991A-11104	Sequence 11104, A
17	48.8	6.6	1044	4	US-09-252-991A-11159	Sequence 11159, A
18	48.8	6.6	1557	4	US-09-252-991A-11093	Sequence 11093, A
19	48	6.5	30001	1	US-08-125-468-1	Sequence 1, Appli
20	48	6.5	30001	2	US-08-474-933-1	Sequence 1, Appli
21	47.8	6.5	30001	1	US-09-252-991A-6608	Sequence 6608, Ap
22	47.8	6.5	558	4	US-09-252-991A-6723	Sequence 6723, Ap
23	47.8	6.5	645	4	US-09-252-991A-6664	Sequence 6664, Ap
24	46.6	6.3	984	4	PCT-US95-04801-4	Sequence 4, Appli
25	46.2	6.2	23673	4	US-09-773-816-1	Sequence 1, Appli
26	45	6.1	876	4	US-09-252-991A-12184	Sequence 12184, A
27	45	6.1	1392	4	US-09-252-991A-12249	Sequence 12249, A
28	45	6.1	1407	4	US-09-252-991A-12361	Sequence 12361, A
29	44.8	6.1	780	4	US-09-252-991A-3741	Sequence 3741, Ap
30	44.8	6.1	960	4	US-09-252-991A-3730	Sequence 3730, Ap
31	44.8	6.1	1089	4	US-09-252-991A-3785	Sequence 3785, Ap
32	44.6	6.0	1656	3	US-09-385-028-14	Sequence 14, Appl
33	44.6	6.0	1656	3	US-09-726-614-14	Sequence 14, Appl
34	44.6	6.0	15079	3	US-09-726-614-1	Sequence 1, Appl
35	44.6	6.0	15079	3	US-09-385-028-1	Sequence 1, Appl
36	44.6	6.0	1512	3	US-09-726-614-6	Sequence 6, Appli
37	44.6	6.1	1512	3	US-08-911-853-6	Sequence 6, Appli
38	44.6	6.0	1512	4	US-09-479-453-6	Sequence 6, Appli
39	44.4	6.0	1512	3	US-09-479-453-6	Sequence 6, Appli
40	44.4	6.0	17612	4	US-08-911-853-29	Sequence 29, Appl
41	44.4	6.0	17612	3	US-09-479-409-29	Sequence 29, Appl
42	44.2	6.0	17612	3	US-09-479-453-29	Sequence 29, Appli
43	44	5.9	921	4	US-09-103-840A-1	Sequence 1, Appli
44	44	5.9	973	4	US-09-252-991A-13020	Sequence 13020, A
45	44	5.9	1491	4	US-09-252-991A-1345	Sequence 1345, Ap

ALIGNMENTS

RESULT 1
US-09-340-620A-48
; Sequence 48, Application US/0934062oA
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bettin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-1240001
; CURRENT APPLICATION NUMBER: US/09/340.620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-340-620A-48

Query Match 100.0%; Score 740; DB 4; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.4e-157;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGTCCGGCTGCAGCGGGGGTGAGCGCGGCGGCAGCGGCGGCGGGGATCCTTGGAGCGGGGGC 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CGCGTCCGGCTGCAGCGGGGGTGAGCGCGGCGGCAGCGGCGGCGGGGATCCTTGGAGCGGGGGC 60

Qy 61 GCGCGCGCCGACGCCCATCCTGGATGCGCGCTGGAGAACTTGACCCCGCGGGGGGGCAAGAAGT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GCGCGCGCCGACGCCCATCCTGGATGCGCGCTGGAGAACTTGACCCCGCGGGGGGGCAAGAAGT 120

Qy 121 TCAAGCTGAAGCTGCTGTCGTGTCGGTGCCGCGTGCGGGCTACGGGCGGCCATCCGGCGGGGGCG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TCAAGCTGAAGCTGCTGTCGTGTCGGTGCCGCGTGCGGGCTACGGGCGGCCATCCGGCGGGGGCG 180

Qy 181 CGCTGCTGTCCATGGACGACGCCTTGGACCTCACCGACACAAGCTGTGGTCAGGCTTCTACCTTGGAGA 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CGCTGCTGTCCATGGACGACGCCTTGGACCTCACCGACACAAGCTGTGGTCAGGCTTCTACCTTGGAGA 240

Qy 241 CCTACGCGCGCCGAGCTTCACCGCTAAGCTGTTGCTGCGCGGACATGGGCCCTGCAGGAGGAGATGCCG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CCTACGCGCGCCGAGCTTCACCGCTAAGCTGTTGCTGCGCGGACATGGGCCCTGCAGGAGGAGATGCCG 300

RESULT 2
US-09-340-620A-50
; Sequence 50, Application US/09340620A
; Patent No. 648293J
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT FILING DATE: 1999-06-28
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match   79.1%;   Score 585;   DB 4;   Length 585;
Best Local Similarity   100.0%;   Pred. No. 8.5e-123;
Matches 585;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
US-09-340-620A-50

Qy   241 CCTACGGCGCGGCGGCGAGCTCACCGCTAACGTGCCGCCGGGGCGCCTGCGGAGATGGCCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTACGGCGCGGCGGCGAGCTCACCGCTAACGTGCCGCCGGGGCGCCTGCGGAGATGGCCG 300

Qy   301 GGCAGCTGCGAGGCGCCACGCACCAGCTGCTTCTGAGCGCCGGATCCAGGGCC   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGCAGCTGCGAGGCGCCACGCACCAGCTGCTTCTGAGCGCCGGATCCAGGGCC   360

Qy   361 CTCCTCAGTCGGCGGAGCCAAGCCAGGCCTTGCACTTTATAGACACCACCACCCGGGCTTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTCCTCAGTCGGCGGAGCCAAGCCAGGCCTTGCACTTTATAGACACCACCACCCGGGCTTA 420

Qy   421 TCGCGAGGGGTCACAAACGTTGAGTGGCTGCTGGCCTGGAGGAAGGTCCTGACGG   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCGCGAGGGGTCACAAACGTTGAGTGGCTGCTGGCCTGGAGGAAGGTCCTGACGG   480

Qy   481 ATGAGACAGTACCAGGCAGTGCCGGCCGAGCCCAAGCCAAGAAGCAAGAAGCTCT   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGACAGTACCAGGCAGTGCCGGCCGAGCCCAAGCCAAGAAGCAAGAAGCTCT   540

Qy   541 TCAGTTTCACACACCAGCCTGGACTGGAGGACTTGCTCTCCTCCCAGGGCCCTAGGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCAGTTTCACACACCAGCCTGGACTGGAGGACTTGCTCTCCTCCCAGGGCCCTAGGG   600

Qy   601 AGTCCCAGTCCTACCTGGTGGAGGACCTGGAGGTCTTGCTCCTTCCCAGCCAACAC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGTCCCAGTCCTACCTGGTGGAGGACCTGGAGGTCTTGCTCCTTCCCAGCCAACAC   660

Qy   661 TCCGGTCAGCCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTTATACACAAT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCCGGTCAGCCCCCTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTTATACACAAT 720

Qy   721 ATACGAAAAGCCAGCTTGAA 740
           ||||||||||||||||||||
Db   721 ATACGAAAAGCCAGCTTGAA 740

RESULT 3
US-09-340-620A-60
; Sequence 60, Application US/09340620A
; Patent No. 648293J
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT FILING DATE: 1999-06-28
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 777
; TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(667)
; ORGANISM: Mus musculus

Query Match   45.4%;   Score 336;   DB 4;   Length 777;
Best Local Similarity   72.9%;   Pred. No. 7.4e-67;
Matches 461;   Conservative   0;   Mismatches 165;   Indels   6;   Gaps   2;
US-09-340-620A-60

Qy   174 CGGGGCGGCCGCTGCGTCGTTCCATGGAGGACCCCTTGGAGACCTCACCGACAAGCTGGTTCAGCTTCTAC 233
           |||||||||||| ||||| || | ||||| |||||| | || || ||||| | ||||||||
Db   121 CGGGGGCGGCCGCTGCGTCTCCATGGAGGACCCTTTGGAGACTTCACCGACGAAGCTGGTTCAGCTTCTAC 180

Qy   234 CTGGAGAACCTACGGCGCGCCGAGCTCACCGGCTTCACCGACATGGGCGCATGAGGAG 293
           ||||||| ||| | || || |||| ||||| ||||| ||| ||||| | ||||||
Db   181 CTGGAGAACCTACGGCGCGCCGAGCTCACCGGCTTCACCGACATGGGCGCATGAGGAG 240

Qy   294 ATGGCCGGGCAGCTGCAGGGGGCGCCACGCACCAGCTGCTTCTGAGCGCCGGGATC 353
           ||||| ||||||||||| || ||||||||||||||||||| |||||| |||||
Db   241 ATGGCCGGGCAGCTGCAGGGGGCGCCACGCACCAGCTGCTTCTGAGCGCCGGGATC 300

Qy   354 CAGGGCCCCTCCTCAGTCGGCGGAGCCAAGCCAGGCCTTGCACTTTATAGACACCAGACACCACCGGGGCT 413
           ||||||| |||||||||| || |||| |||||||| |||||||||||||||| || |||| ||||||
Db   301 CAGGGCCCCTCCTCAGTCGGCGGAGCCAAGCCAGGCCTTGCACTTTATAGACACCACCACCCGGGCT 360

Qy   414 GCGTTTATCGCGAGGGGTCACAAACGTTGAGTGGCTGCTGGAGGTC 473
           ||| || ||||| |||||||||||||||||||||||||| ||||||
Db   361 GCGCTTTATCGCGAGGGGTCACAAACGTTGAGTGGCTGCTGGCCTGGAGGTC 420

Qy   474 CTGACGGGATGAGACAGTACCAGGCAGTGCCGGCCGAGCCCAAGCCAAGAAGCAAGATGCGG 533
           |||||||| || ||||| ||| ||| || |||||| || |||||| || || |||||
Db   421 CTGACGGGATGAGACAGTACCAGGCAGTGCCGGCCGAGCCCAAGCCAAGAAGCAAGAAGCTCT 480

Qy   534 AAGCTCTTCAGTTTCACACACCAGCCTGGAACTGGAGGACTTGCTCTCCTCCCAGGGCC 593
           |||||||||||||||||| | |||||||| |||||||||||||| |||||||||||||
Db   481 AAGCTCTTCAGTTTCACACACCAGCCTGGACTGGAGGACTTGCTCTCCTCCCAGGGCC 540

Qy   594 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGGGAGC 638
           |||||| |||||||||||||||||||||| ||||||||||||||
Db   541 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGGGGAGC 585

RESULT 2
US-09-340-620A-50
; Sequence 50, Application US/09340620A
; Patent No. 648293J
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE

Query Match   79.1%;   Score 585;   DB 4;   Length 585;
Best Local Similarity   100.0%;   Pred. No. 8.5e-123;
Matches 585;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
US-09-340-620A-50

Qy   54   ATGGGGCGCGGCGGCGACCGCATCCTGGATGCGCTGGAGGAACCTGACCCGCGACGGAGCTC 113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   ATGGGGCGCGGCGGCGACCGCATCCTGGATGCGCTGGAGGAACCTGACCCGCGACGGAGCTC 60

Qy   114 AAGAAGTTCAAGCTGAAGCTGCTCGTGTCGGTGCCCGCTGCCGGAGGGCTACGGGGGGCCATCCCG 173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   AAGAAGTTCAAGCTGAAGCTGCTCGTGTCGGTGCCCGCTGCCGGAGGGCTACGGGGGGCCATCCCG 120

Qy   78   CTGGAATGCGCTGGAGAGAACCTGACCCGCGACGGAGGAGCTCAAGAGAGTTCAAGCTGAAGCTGCTG 137
           || ||| || |||||| |||||| || ||||| ||| | || ||| |||||||||||| |||||
Db   53   GAGTAAAAGGTGACCGCGCGCCACCCCACAGCGGAGCTCAAGAGAGATCCATC 112

```
RESULT 4
US-09-340-620A-62
; Sequence 62, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
;   APPLICANT: Bertin, John
;   TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;   FILE REFERENCE: 07334-124001
;   CURRENT APPLICATION NUMBER: US/09/340,620A
;   CURRENT FILING DATE: 1999-06-28
;   PRIOR APPLICATION NUMBER: US 09/245,281
;   PRIOR FILING DATE: 1999-02-05
;   PRIOR APPLICATION NUMBER: US 09/207,359
;   PRIOR FILING DATE: 1998-12-08
;   PRIOR APPLICATION NUMBER: US 09/099,041
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: US 09/019,942
;   PRIOR FILING DATE: 1998-02-06
;   NUMBER OF SEQ ID NOS: 71
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 62
;   LENGTH: 579
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-340-620A-62

  Query Match             42.7%;  Score 316.2;  DB 4;  Length 579;
  Best Local Similarity   73.7%;  Pred. No. 1.9e-62;
  Matches   431;  Conservative   0;  Mismatches  148;  Indels    6;  Gaps    2;

Qy     54 ATGGGGGCGGCGCGCGGACGCCATCCTCTGGATGGCGGTCTTGAAAAACTTGTCACCGCGCGAGGAGCTC 113
          ||||||||||| ||  ||||||||||||||||||||||||||||||||||||  |||||| |||||  |||
Db      1 ATGGGGGCGGGGCCACGAGATGCCATCCTCTTGGACGCTCTTGAAAAACTTGTCACCGCGCGAGGAGCTC 60

Qy    114 AAGAAGTTCAAGCTGAAGCTCGCTTGTGTGGGTGCCGCTGCGCGAGGGCTACGGGCGGGCATCCCG 173
          ... 
Db     61 AAAAAGTTCAAGATGAAGCTGCTGAC... 120
```

```
RESULT 5
US-09-252-991A-2958
; Sequence 2958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
;   SEQ ID NO 2958
;   LENGTH: 909
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2958

  Query Match             7.2%;  Score 53;  DB 4;  Length 909;
  Best Local Similarity   49.1%;  Pred. No. 0.0027;
  Matches   140;  Conservative   0;  Mismatches  145;  Indels    0;  Gaps    0;
```

RESULT 6
US-09-252-991A-3146
Sequence 3146, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3146
;   LENGTH: 1218
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3146

Query Match   49.1%;   Score 53;   DB 4;   Length 1218;
Best Local Similarity 49.1%;   Pred. No. 0.0029;
Matches 140; Conservative   0; Mismatches 145; Indels   0; Gaps   0;

Qy   63 GCGCGCGACGCCATCCCTGGATGCGCTGGAGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTC 122
               |||||||| ||||
Db   925 GCGCGCGCTCTCACCACCGAATACGAAGGCGGCAAGGCGGTTCAAGGAAGTCGCGCGGTTC 984

Qy   123 AAGCTGAAGCTGCTGTTCGGTGTCCGCTGCCGGAGGAGCTACGGGGGCGCGCG 182
               |||| ||||
Db   985 CGCTCCTACTCGTCGTTCGGTGTCCGGTGCCGGAGGAGCTACGGGGGCAAC 1044

Qy   183 CTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTTCAGCTTCTTACCTGGATACC 242
               |||| ||||
Db   1045 GACCGCTACCAGAACGCCCTGGACAGCGCGCCAACCCCGAGCGCTTCATGCAGGAAACTG 1104

Qy   243 TACGGCGCGCCGAGCTCACCGCTAACGTGCTCGCGGACATGGGCCTGCAGGAGATGGCCGGG 302
               |||| ||||
Db   1105 CAGCGCGCCGGGCGTACGCCACCAGGCCGATCCGCCCCGCAAGGTNGGCGCGAGATCGCCAGA 1164

Qy   303 CAGCTGCAGGCGGGCGACGACCAGGGCTCTGGACCGCCGCCAGCT 347
               ||||
Db   1165 CAGATGCAGACCTACCAGGCCGTCCGCCCGGCACGCGCGCCT 1209

RESULT 7
US-09-252-991A-2822/c
Sequence 2822, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2822
;   LENGTH: 1380
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2822

Query Match   49.1%;   Score 53;   DB 4;   Length 1380;
Best Local Similarity 49.1%;   Pred. No. 0.003;
Matches 140; Conservative   0; Mismatches 145; Indels   0; Gaps   0;

Qy   63 GCGCGCGACGCCATCCCTGGATGCGCTGGAGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTC 122
               |||||||| ||||
Db   321 GCGCGCGCTCTCACCACCGAATACGAAGGCGGCAAGGCGGTTCAAGGAAGTCGCGCGGTTC 262

Qy   123 AAGCTGAAGCTGCTGTTCGGTGTCCGCTGCCGGAGGAGCTACGGGGGCGCGCG 182
               |||| ||||
Db   261 CGCTCCTACTCGTCGTTCGGTGTCCGGTGCCGGAGGAGCTTCCACGACTTCCACGACAAC 202

Qy   183 CTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTTCAGCTTCTTACCTGGATACC 242
               |||| ||||
Db   201 GACCGCTACCAGAACGCCCTGGACAGCGCGCCAACCCCGAGCGCTTCATGCAGGAAACTG 142

Qy   243 TACGGCGCGCCGAGCTCACCGCTAACGTGCTCGCGGACATGGGCCTGCAGGAGATGGCCGGG 302
               |||| ||||
Db   141 CAGCGCGCCGGGCGTACGCCACCAGGCCGATCCGCCCCGCAAGGTNGGCGCGAGATCGCCAGA 82

Qy   303 CAGCTGCAGGCGGGCGACGACCAGGGCTCTGGACCGCCGCCAGCT 347
               ||||
Db   81 CAGATGCAGACCTACCAGGCCGTCCGCCCGGCACGCGCGCCT 37

RESULT 8
US-09-252-991A-3050
Sequence 3050, Application US/09252991A
Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3050
;   LENGTH: 1500
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3050

Query Match   49.1%;   Score 53;   DB 4;   Length 1500;
Best Local Similarity 49.1%;   Pred. No. 0.003;
Matches 140; Conservative   0; Mismatches 145; Indels 1500; Gaps   0;

Qy   63 GCGCGCGACGCCATCCCTGGATGCGCTGGAGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTC 122
               |||||||| ||||
Db   920 GCGCGCGCTCTCACCACCGAATACGAAGGCGGCAAGGCGGTTCAAGGAAGTCGCGCGGTTC 979

Qy   123 AAGCTGAAGCTGCTGTTCGGTGTCCGCTGCCGGAGGAGCTACGGGGGCGCGCG 182
               |||| ||||
Db   980 CGCTCCTACTCGTCGTTCGGTGTCCGGTGCCGGAGGAGCTACGGGGGCAAC 1039

Qy   183 CTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTTCAGCTTCTTACCTGGATACC 242
               |||| ||||
Db   1040 GACCGCTACCAGAACGCCCTGGACAGCGCGCCAACCCCGAGCGCTTCATGCAGGAAACTG 1099

Qy   243 TACGGCGCGCCGAGCTCACCGCTAACGTGCTCGCGGACATGGGCCTGCAGGAGATGGCCGGG 302
               |||| ||||
Db   1100 CAGCGCGCCGGGCGTACGCCACCAGGCCGATCCGCCCCGCAAGGTNGGCGCGAGATCGCCAGA 1159

RESULT 9
US-09-252-991A-3097
; Sequence 3097, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3097
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3097

Query Match   6.8%;   Score 50.6;   DB 4;   Length 2079;
Best Local Similarity   50.5%;   Pred. No. 0.011;
Matches 150; Conservative   0; Mismatches 144; Indels   3; Gaps   1;

Qy   7 CGGCTGACGCAGGGGGTGAGCGGCGCCAGGGGGGGGATCCTTGCGAGCCATGGGGCGCCGC 66
            ||| || |||| ||| || ||| || || |||| | | |||| || ||||| || ||
Db   401 CGGCCGCAGCGCAGTTCGACGCGGATCGCCACCGGGGGTCACGCGGTGAACCTTGGAGGGGCGCGC 460

Qy   67 GCGACGCCATCCTGGATGCGCTGGAGAAACCTGACCCCGCGAGAGCTCAAGAAGTTCAAGC 126
            || ||||| ||| ||| || ||| ||| | ||| | || || ||||| |||||| || |
Db   461 GCCAGGTCCACGAGCGGTTGCTCGGAGGAGGGCGAAGAACGATCCGGCCGCGGGGCCCAGCC 520

Qy   127 TGAAGCTGC----TGTCGGTCCCGCTGCGCGCAGGGGCTACGGGGCGCATCCCCGCCGGCGCGC 183
            || |||| || || || || || || || || | | | || | | ||| || |
Db   521 TGCTCGCTGCGCCTTCACGGAGTTGCAGACGCGGGTGATCGCCTTACTGACCGCGGGTCGGCC 580

Qy   184 TGCTCGTTCCATTGGACGGCCTTGGACCTCACCGACAAGCTGGTGTCAGCTTCTACCTGGAGACCT 243
            || || ||| |||| || || ||| || ||| || | | ||||| || |||||||| ||
Db   581 TGCTCGTCGACGATATTCTCAAAGACGCGCAGCAACCTGGTTACCCAGGGGGGGGCAGGGCG 640

Qy   244 ACGGCGCCGAGCTCACCGCTAACGTCCTCGCCGACATGGGGCCCTGCAGGAGAGGATGGCCG 300
            |||| ||| | |||| | || | ||| ||| ||||| | ||| | ||| || ||||
Db   641 ACGGCCTGAATCGCTTCCGCGCGGGTGTTCGGCACCCTGCCGCCTGCCGGAGTGGTCGCG 697

RESULT 10
US-09-144-085-3
; Sequence 3, Application US/0914085
; Patent No. 6280999
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Claes
; APPLICANT: Betlach, Mary C.
; APPLICANT: Ashley, Gary
; APPLICANT: Julien, Bryan
; APPLICANT: Ziermann, Rainer
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 30062-20020.20
; CURRENT APPLICATION NUMBER: US/09/144,085
; CURRENT FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010,809
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3

Query Match   6.7%;   Score 49.8;   DB 3;   Length 33529;
Best Local Similarity   51.6%;   Pred. No. 0.032;
Matches 114; Conservative   0; Mismatches 107; Indels   0; Gaps   0;

Qy   90 GAGAACCTGACCGCCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGCGTGGTGCCGTG 149
            ||||| ||||| || |||| || ||||||||| ||||| |||| || | | | |||
Db   14083 GAGACGGTGACGGCGTGGCCGTGGCCGACGTGTCGAAGGGGAGGAGGTCGGCGGGTGCTG 14142

Qy   150 CGCGGAGGGCCTACGGGGCGGCGGCCCATCCCGCGGGGGGGGGCCGCTTGGACCTC 209
            || ||||| || | ||| || ||| || || ||| |||| || || |||||
Db   14143 GCCGGCATCGACGGCGGCGGGCGGGCCGGGCCCTGAGCGCGGGTGTGCACCTGGCGCGGCTCGAC 14202

Qy   210 ACCGACAAGCTGGTTCAGCTTCTACCTGGAGACCTACGGGCGCCGGAGCTTCACCGGCTAACGTG 269
            ||||| || | || ||| || ||| |||| ||||| || |||| | ||||| || |
Db   14203 GACGGCGGTTCGTCGCGGCCCAGCAGCGCCCTTCGCGGCGGGTGCTTGGGCGCCGAAGGTG 14262

Qy   270 CTGCGCGCACATGGGCCTGCGCAGGAGATGGCCGGGCAGCTGCA 310
            ||| ||| |||| | || || ||| || || || || ||||
Db   14263 GACGGGGCGGCTGCCACCTGCGCAGGAGCTGAACGGCGGGAGCTGGA 14303

RESULT 11
US-09-428-517-1/c
; Sequence 1, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 50937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-428-517-1

Query Match   6.6%;   Score 49.2;   DB 3;   Length 50937;
Best Local Similarity   47.7%;   Pred. No. 0.047;
Matches 144; Conservative   0; Mismatches 158; Indels   0; Gaps   0;

Qy   86 GCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCC 145
            |||||| |||||| | | ||| || ||||||||| |||||||| || | | | | |||
Db   14083 GCGGGAGCGGCTGGGCGGCAGCGGCGGCTGGGCCGCGGCTGATCCACTACCGGTGCCGGTGCA 1691

Qy   146 GCTGCGCGAGGGCTACGGGCGCCGCATCCCGCGGGGGCGCGCCGTGCGTCCATGGACGGCCTTGGA 205
            ||||| ||||| || | ||| || || |||||||| | || || |||| | ||||| ||
Db   1690 CCGGTCCGAGCGGGGTACCCCGGCAGCCGCCGCGGCGACCAGCCCGTGCGGGAGGCG 1631

Qy   206 CCTCACCGGACAAGCTGGTTCAGGTTCTACCTGGAGACCTACGGGCGCCGGAGCTCACCGGCTAA 265
            || |||| |||| ||||| ||| || ||| |||| ||| | |||||| | || |
Db   1630 GCTCGCCCGGGAGGGTGCTTGACGCCTGCCCATAGGCCGCACTTGTCCGGACAGACGCCGGGTCAA 1571

Qy   266 CGTGCTCGCGCGACATGGGCGTTCGGCGGGGAGGATGGCCGAGCGTGCAGGCGGCGGCCACGCACCA 325
            || | | | | | |||| | || | |||| ||||| | ||||| || | ||||| | |
Db   1570 GGCGGGTGATCGAGGCCGGTTACGGGGGGGGCGGCGGCTGCCGGAGAGTCCGGGGGGGCGCG 1511

```
Qy    326  GGGCTCTTGAGCCCCGGATCATCAGGCCGCCCCCTCCTTCAGTTCGGCGAGCCAAGCCACAGG  385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1510  TCGCCCCCGGTACCCGTCGGGTACGGGGGGGCCGACTTCACCGCCGCGGGGCCGGGCGGC  1451

Qy    386  CC  387
           ||
Db   1450  CC  1449
```

RESULT 12
US-09-252-991A-1107/c
; Sequence 1107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
;   SEQ ID NO 1107
;   LENGTH: 2610
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1107

  Query Match   6.6%;   Score 49; DB 4; Length 2610;
  Best Local Similarity 49.8%;   Pred. No. 0.027;
  Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
Qy     62  CGGCGGCGGACGGCCATCCTTGGATNGGGCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTT  121
           |||||  ||| ||||||        ||||| |||||||||||||| ||||||| |||| ||||
Db   1942  CGAGCGGCGGCCATTCCGCTGCACGCGCCCGAGGACGTGAACCATCCGCTGTTGGCGTGGAGCG  1883

Qy    122  CAAGCTGAAGCTGCTGTCGTGCCGGCTGCCGCGAGGAGGGCTACGGGGCGCATCCGCGCGGGGCGC  181
           |||||||| || || |||||| |||||| ||| ||| |||||| |||| |||| ||||| ||||
Db   1882  CATCCGCCAACTGGCCCCGGACTTCCTGTTCTCTTCTACTACCGCCGCCTGCTCGGGCGC  1823

Qy    182  GCTGCTGTCCATGGACGCCTTGGACCTCACCGACACAAGCTGGTCAGCTTCTTACCTGGAGAC  241
           ||||| |||||| |||| | |||  || | ||||| | ||  ||||| || |||| ||||||
Db   1822  CGAGCTGCTTGCCTTGCGCCGCCGCGCCTACAACCTGCAGGTTCGCTGCTGCGGCGCG  1763

Qy    242  CTACGGCGCGGCCGAGCTCACCGCCTAACGTGTCGTGCGCGGACATGGGCCTGCGCAGGAGATGGCCGG  301
           |||||||||||||||| |||||||||||||||||||||||||  || |||||| |||||| |||
Db   1762  CTACCGCGGACGAGGCCCGGCGAACTGGGTGCTGGTCAACGGGCGAAACGCAGACCGGGGGT  1703

Qy    302  GCAGCTGCA  310
           |||||||||
Db   1702  GACCCTGCA  1694
```

RESULT 13
US-09-252-991A-977
; Sequence 977, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142

; LENGTH: 3114
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-977

  Query Match   6.6%;   Score 49; DB 4; Length 3114;
  Best Local Similarity 49.8%;   Pred. No. 0.028;
  Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
Qy     62  CGGCGGCGGACGGCCATCCTTGGATNGGGCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTT  121
           |||||  ||| ||||||        ||||| |||||||||||||| ||||||| |||| ||||
Db   1281  CGAGCGGCGGCCATTCCGCTGCACGCGCCCGAGGACGTGAACCATCCGCTGTTGGCGTGGAGCG  1340

Qy    122  CAAGCTGAAGCTGCTGTCGTGCCGGCTGCCGCGAGGAGGGCTACGGGGCGCATCCGCGCGGGGCGC  181
           |||||||| || || |||||| |||||| ||| ||| |||||| |||| |||| ||||| ||||
Db   1341  CATCCGCCAACTGGCCCCGGACTTCCTGTTCTCTTCTACTACCGCCGCCTGCTCGGGCGC  1400

Qy    182  GCTGCTGTCCATGGACGCCTTGGACCTCACCGACACAAGCTGGTCAGCTTCTTACCTGGAGAC  241
           ||||| |||||| |||| | |||  || | ||||| | ||  ||||| || |||| ||||||
Db   1401  CGAGCTGCTTGCCTTGCGCCGCCGCGCCTACAACCTGCAGGTTCGCTGCTGCGGCGCG  1460

Qy    242  CTACGGCGCGGCCGAGCTCACCGCCTAACGTGTCGTGCGCGGACATGGGCCTGCGCAGGAGATGGCCGG  301
           |||||||||||||||| |||||||||||||||||||||||||  || |||||| |||||| |||
Db   1461  CTACCGCGGACGAGGCCCGGCGAACTGGGTGCTGGTCAACGGGCGAAACGCAGACCGGGGGT  1520

Qy    302  GCAGCTGCA  310
           |||||||||
Db   1521  GACCCTGCA  1529
```

RESULT 14
US-09-252-991A-937
; Sequence 937, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;   APPLICANT: Marc J. Rubenfield et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;   TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196.136
;   CURRENT APPLICATION NUMBER: US/09/252,991A
;   CURRENT FILING DATE: 1999-02-18
;   PRIOR APPLICATION NUMBER: US 60/074,788
;   PRIOR FILING DATE: 1998-02-18
;   PRIOR APPLICATION NUMBER: US 60/094,190
;   PRIOR FILING DATE: 1998-07-27
;   NUMBER OF SEQ ID NOS: 33142
;   SEQ ID NO 937
;   LENGTH: 3195
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-937

  Query Match   6.6%;   Score 49; DB 4; Length 3195;
  Best Local Similarity 49.8%;   Pred. No. 0.028;
  Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
Qy     62  CGGCGGCGGACGGCCATCCTTGGATNGGGCTGGAGAACCTGACCGCCCGAGGAGCTCAAGAAGTT  121
           |||||  ||| ||||||        ||||| |||||||||||||| ||||||| |||| ||||
Db    478  CGAGCGGCGGCCATTCCGCTGCACGCGCCCGAGGACGTGAACCATCCGCTGTTGGCGTGGAGCG  537

Qy    122  CAAGCTGAAGCTGCTGTCGTGCCGGCTGCCGCGAGGAGGGCTACGGGGCGCATCCGCGCGGGGCGC  181
           |||||||| || || |||||| |||||| ||| ||| |||||| |||| |||| ||||| ||||
Db    538  CATCCGCCAACTGGCCCCGGACTTCCTGTTCTCTTCTACTACCGCCGCCTGCTCGGGCGC  597

Qy    182  GCTGCTGTCCATGGACGCCTTGGACCTCACCGACACAAGCTGGTCAGCTTCTTACCTGGAGAC  241
           ||||| |||||| |||| | |||  || | ||||| | ||  ||||| || |||| ||||||
Db    598  CGAGCTGCTTGCCTTGCGCCGCCGCGCCTACAACCTGCAGGTTCGCTGCTGCGGCGCG  657

Qy    242  CTACGGCGCGGCCGAGCTCACCGCCTAACGTGTCGTGCGCGGACATGGGCCTGCGCAGGAGATGGCCGG  301
           |||||||||||||||| |||||||||||||||||||||||||  || |||||| |||||| |||
Db    658  CTACCGCGGACGAGGCCCGGCGAACTGGGTGCTGGTCAACGGGCGAAACGCAGACCGGGGGT  717

Qy    302  GCAGCTGCA  310
```

Db   718 GACCCTGCA 726
            |||||
            |

RESULT 15
US-09-252-991A-10961/c
; Sequence 10961, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10961
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10961

Query Match   6.6%;   Score 48.8; DB 4; Length 468;
Best Local Similarity   47.4%;   Pred. No. 0.02; Indels   0; Gaps   0;
Matches 146; Conservative   0; Mismatches 162;   Indels   0; Gaps   0;

Qy   1   CGCGTCCGGCTGCACGCGGGGGTGAGCGGCGGCGCCGGGGATCCTGGAGCCATGGGGC   60
          ||||||||||||||||||||||||||||||||||
Db   325 CGCTGCCGGTCGACACGGGTTCGACGGAGTTCGAGGCCCTGTGCTCGAGGAGCAGGTCC   266

Qy   61 GCGCGGCGGACGCCATCCTGGAATGCGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAAGT   120
          |||||||||
Db   265 GCGGATGGCGCGCCTGGTTCCGCCGGCGGGCCCTTGCTCGACGAAACCGACGACGTTGC   206

Qy   121 TCAAGCTGAAGCTGCTGTGTCGGTGCCGCGCCGGAGGGCCATCCCGCGGGGGGG   180
          |||||||||
Db   205 GCGAGCTGCCGGTCGAGCCACGGGCCCACGAAAGTCATCGAGTTGCCGGTGCCGAGCGCCC   146

Qy   181 CGCTGCTGTCATGGACGCGCTTGGACCTCACCGACAAGCTGGTCAGCTTCTTACCTGGAGA   240
          |||||||
Db   145 CGGGCGAGCGTCGGGCAGCCGCGGCCTCGACCTGCCGCGCTCTCGACGACGGGTCGTCCGG   86

Qy   241 CCTACGGCGCGCCGAGCTCACCGCTAACGTGCGCGGACATGGGCGTTGCAGGAGATGGCCG   300
          |||||||||
Db   85 CCTGGGCGCGGGGGAGCGCCCTTCGAGTGGCCTGCTCGTTCGACGTGCGCCGGCGGCGGGCCTGGGCGG   26

Qy   301 GGCCAGCTG   308
          |||||||
Db   25 TGCCGGCTG   18

Search completed: January 29, 2004, 13:11:29
Job time : 41.6087 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 29, 2004, 03:03:41

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcgggg..........atacgaaaagccagctttgaa 740

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 182278265 residues

Total number of hits satisfying chosen parameters: 4865878

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :
Published_Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Search time 235.019 Seconds
(without alignments)
11475.528 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	740	100.0	740	9	US-09-728-721-48	Sequence 48, Appl
2	740	100.0	740	10	US-09-996-617-7	Sequence 7, Appli
c 3	740	100.0	740	10	US-09-728-721-60	Sequence 10, Appl
c 4	740	100.0	740	10	US-09-841-879B-4	Sequence 4, Appli
c 5	740	100.0	740	15	US-09-841-879B-19	Sequence 19, Appl
c 6	740	100.0	740	15	US-10-295-981-48	Sequence 48, Appl
7	733	99.1	745	12	US-10-240-145-12	Sequence 12, Appl
8	733	99.1	779	12	US-10-131-410-21	Sequence 21, Appl
c 9	721.2	97.5	806	15	US-10-106-698-1144	Sequence 1144, Ap
c 10	697.6	94.3	811	9	US-09-925-301-278	Sequence 278, App
11	585	79.1	585	10	US-09-728-721-50	Sequence 50, Appl
12	585	79.1	585	10	US-09-996-617-9	Sequence 9, Appl
13	585	79.1	585	10	US-09-841-879B-6	Sequence 6, Appl
14	585	79.1	585	15	US-10-295-981-50	Sequence 50, Appl
15	531	71.8	639	15	US-10-106-698-1145	Sequence 1145, Ap
16	384.6	52.0	432	9	US-09-925-299-648	Sequence 648, App
17	384.6	52.0	432	11	US-09-728-721-60	Sequence 648,
18	336	45.4	777	9	US-09-728-721-60	Sequence 60,
19	336	45.4	777	10	US-09-841-879B-1	Sequence 1, Appli
20	336	45.4	777	15	US-09-841-879B-18	Sequence 18, Appl
21	336	45.4	777	15	US-09-728-721-62	Sequence 62, Appl
22	316.2	42.7	579	9	US-09-728-721-62	Sequence 62, Appl
23	316.2	42.7	579	12	US-10-295-981-62	Sequence 62, Appl
c 24	316.2	41.9	321	12	US-09-880-107-1066	Sequence 1066, Ap
c 25	293.8	39.7	579	10	US-09-841-879B-3	Sequence 3, Appli
26	244.6	33.1	551	12	US-10-191-803-602	Sequence 602, Appl
27	169.4	22.9	270	15	US-10-295-621-27	Sequence 27, Appli
28	169.4	22.9	270	10	US-09-965-621-27	Sequence 27, Appl
29	151.2	20.4	437	12	US-10-242-674-985	Sequence 985, App
30	146.4	19.8	174	12	US-10-242-535A-45706	Sequence 45706, A
31	115	15.5	4200	10	US-09-388-221-3	Sequence 3, Appli
32	115	15.5	4287	10	US-09-931-071-5	Sequence 5, Appli
33	115	15.5	4287	10	US-09-388-221-5	Sequence 5, Appli
34	115	15.5	4332	10	US-09-388-221-5	Sequence 5, Appli
35	115	15.5	4422	10	US-09-388-221-1	Sequence 1, Appli
36	115	15.5	5100	11	US-09-956-712-12	Sequence 12, Appli
37	115	15.5	5444	10	US-09-956-617-1	Sequence 1, Appli
38	115	15.5	5444	11	US-09-931-071-1	Sequence 1, Appli
39	115	15.5	6531	11	US-09-207-359	Sequence 1, Appli
40	115	15.5	6531	10	US-09-895-298-22	Sequence 22, Appl
41	111.4	15.1	2657	11	US-09-956-712-11	Sequence 11, Appli
42	110.6	14.9	96649	11	US-09-956-712-10	Sequence 10, Appli
43	98.6	13.3	180	12	US-10-240-145-98	Sequence 98, Appl
44	74	10.0	394	9	US-09-864-761-21179	Sequence 21179, A
45	70.6	9.5	2708	12	US-10-407-866-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1
US-09-728-721-48
; Sequence 48, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-728-721-48

Query Match 100.0%; Score 740; DB 9; Length 740;
Best Local Similarity 100.0%;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 CGCGTCCGGCTGCAGGGGGTTGCAGCGGCGGCGGCCGCCGGGGGATCCTGGAGCCATGGGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGCGTCCGGCTGCAGGGGGTTGCAGCGGCGGCGGCCGCCGGGGGATCCTGGAGCCATGGGC 60

Qy  61 GCGCGCGGCGACGCCATCCTGATGCGCCTGAGAACCTGACCGCCGAGGAGTCAAGAAGT 120
```

```
Qy    61  GCCGCGCGGGACGGCCATCCTGGATCCGCTGGAGAACCTGACCCGGCGGGATCCAAGAAGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TCAAGCTGAAGCTGCTGTCGGTGCGCTGCGGGAGGGGTACGGGCCGCATCCCGCGGGGCG 180
```

(sequence alignment block — Qy/Db rows continuing through position 740)

```
Qy   721  ATACGAAAAGCCAGCTTGAA 740
          ||||||||||||||||||||
Db   721  ATACGAAAAGCCAGCTTGAA 740
```

RESULT 2
US-09-996-617-7
; Sequence 7, Application US/0996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (54)...(638)
US-09-996-617-7
  Query Match   100.0%; Score 740; DB 10; Length 740;
  Best Local Similarity 100.0%; Pred. No. 1.1e-196;
  Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 CGCGTCCGGCTGCAGCGCGGGGTGACGGCCGGCAGCGGCCGCGGGGATCCGTGGAGCAAGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CGCGTCCGGCTGCAGCGCGGGGTGACGGCCGGCAGCGGCCGCGGGGATCCGTGGAGCAAGC 60
```

(sequence alignment block continues through position 740)

```
Qy   721  ATACGAAAAGCCAGCTTGAA 740
          ||||||||||||||||||||
Db   721  ATACGAAAAGCCAGCTTGAA 740
```

RESULT 3
US-09-996-617-10/c
; Sequence 10, Application US/0996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
```

; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-10

Query Match   100.0%; Score 740; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGCGTCCGGCTGCACGGGGGTGAGCGGGCGCAGCGGCGCCGGGGGATCCTTGGAGCCATGGGGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    740 CGCGTCCGGCTGCACGGGGGTGAGCGGGCGCAGCGGCGCCGGGGGATCCTTGGAGCCATGGGGC 681

Qy     61 GCGCGCCGACGCCATCCTGGATGGCCTGGAGAACCTGACCGCGGGAGCTCAAGAAGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    680 GCGCGCCGACGCCATCCTGGATGGCCTGGAGAACCTGACCGCGGGAGCTCAAGAAGT 621

Qy    121 TCAAGCTGAAGCTGCTGTCGGTGCCCGCTGCCCGCGAGGGCCATCCCGCGCGGGGGCG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    620 TCAAGCTGAAGCTGCTGTCGGTGCCCGCTGCCCGCGAGGGCCATCCCGCGCGGGGGCG 561

Qy    181 CGCTGCTGTCCATGGACGCCTTGGACCTCACCGACCAAGCTTCTACCTGGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 CGCTGCTGTCCATGGACGCCTTGGACCTCACCGACCAAGCTTCTACCTGGAGA 501

Qy    241 CCTACGGCGCCGAGCTCACCGGCTAAGGTGTGCGGCGACATGGGCGCTACGGCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500 CCTACGGCGCCGAGCTCACCGGCTAAGGTGTGCGGCGACATGGGCGCTACGGCCG 441

Qy    301 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGACCGCGGCGAATCCAGGCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    440 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGACCGCGGCGAATCCAGGCCC 381

Qy    361 CTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCCGGCGCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    380 CTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCCGGCGCTTA 321

Qy    421 TCGCGAGGGGTCACAAACGTTGAGTGGCTCGTGTACGGGGAAGGTCCTGACGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    320 TCGCGAGGGGTCACAAACGTTGAGTGGCTCGTGTACGGGGAAGGTCCTGACGG 261

Qy    481 ATGAGCAGTACCAGGCGCAGTGCGGGCGCCCAGCCAAGCAAGATGCGGGAAGCTCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 ATGAGCAGTACCAGGCGCAGTGCGGGCGCCCAGCCAAGCAAGATGCGGGAAGCTCT 201

Qy    541 TCAGTTTCACCAGCAGCCTGGAACTGGACCTTGCTCCTCCAGGCGCCTAAGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 TCAGTTTCACCAGCAGCCTGGAACTGGACCTTGCTCCTCCAGGCGCCTAAGGG 141

Qy    601 AGTCCCAGTCCTACCTGGGTGGAGGAGCTGAGGCGTCCTTCCCAGCGAAACAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 AGTCCCAGTCCTACCTGGGTGGAGGAGCTGAGGCGTCCTTCCCAGCGAAACAC 81

Qy    661 TCCGGTGCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTTATACACAAT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 TCCGGTGCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTTATACACAAT 21

Qy    721 ATACGAAAAGCCAGCTTGAA 740
          ||||||||||||||||||||
Db     20 ATACGAAAAGCCAGCTTGAA 1
```

RESULT 4
US-09-841-879B-4
; Sequence 4, Application US/0984187 9B
; Patent No. US20020142979A1

; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-841-879B-4

Query Match   100.0%; Score 740; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGCGTCCGGCTGCACGGGGGTGAGCGGGCGCAGCGGCGCCGGGGGATCCTTGGAGCCATGGGGC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGCGTCCGGCTGCACGGGGGTGAGCGGGCGCAGCGGCGCCGGGGGATCCTTGGAGCCATGGGGC 60

Qy     61 GCGCGCCGACGCCATCCTGGATGGCCTGGAGAACCTGACCGCGGGAGCTCAAGAAGT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCGCGCCGACGCCATCCTGGATGGCCTGGAGAACCTGACCGCGGGAGCTCAAGAAGT 120

Qy    121 TCAAGCTGAAGCTGCTGTCGGTGCCCGCTGCCCGCGAGGGCCATCCCGCGCGGGGGCG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAGCTGAAGCTGCTGTCGGTGCCCGCTGCCCGCGAGGGCCATCCCGCGCGGGGGCG 180

Qy    181 CGCTGCTGTCCATGGACGCCTTGGACCTCACCGACCAAGCTTCTACCTGGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGCTGCTGTCCATGGACGCCTTGGACCTCACCGACCAAGCTTCTACCTGGAGA 240

Qy    241 CCTACGGCGCCGAGCTCACCGGCTAAGGTGTGCGGCGACATGGGCGCTACGGCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTACGGCGCCGAGCTCACCGGCTAAGGTGTGCGGCGACATGGGCGCTACGGCCG 300

Qy    301 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGACCGCGGCGAATCCAGGCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGCAGCTGCAGGCGGCCACGCACCAGGGCTCTGGAGACCGCGGCGAATCCAGGCCC 360

Qy    361 CTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCCGGCGCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCACTTTATAGACCAGCACCCGGCGCTTA 420

Qy    421 TCGCGAGGGGTCACAAACGTTGAGTGGCTCGTGTACGGGGAAGGTCCTGACGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCGCGAGGGGTCACAAACGTTGAGTGGCTCGTGTACGGGGAAGGTCCTGACGG 480

Qy    481 ATGAGCAGTACCAGGCGCAGTGCGGGCGCCCAGCCAAGCAAGATGCGGGAAGCTCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCAGTACCAGGCGCAGTGCGGGCGCCCAGCCAAGCAAGATGCGGGAAGCTCT 540

Qy    541 TCAGTTTCACCAGCAGCCTGGAACTGGACCTTGCTCCTCCAGGCGCCTAAGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCAGTTTCACCAGCAGCCTGGAACTGGACCTTGCTCCTCCAGGCGCCTAAGGG 600

Qy    601 AGTCCCAGTCCTACCTGGGTGGAGGAGCTGAGGCGTCCTTCCCAGCGAAACAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGTCCCAGTCCTACCTGGGTGGAGGAGCTGAGGCGTCCTTCCCAGCGAAACAC 660

Qy    661 TCCGGTGCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTTATACACAAT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCCGGTGCAGCCCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTTATACACAAT 720
```

RESULT 5
US-09-841-879B-19/c
Sequence 19, Application US/09841879B
Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 19
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-19

Query Match   100.0%; Score 740; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGGGTCCGGCTGCAGCGGCGGGTGAGCGGCCGCAGCGGCCGGGATCCTTGAGCCATGGGGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    740 CGGGTCCGGCTGCAGCGGCGGGTGAGCGGCCGCAGCGGCCGGGATCCTTGAGCCATGGGGC 681

Qy     61 GGCGCGCGACCGCCATCCTCGGATGCGCTGGAGAACCTGACCCGCCGGGGAGCTCAAGAAGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    680 GGCGCGCGACCGCCATCCTCGGATGCGCTGGAGAACCTGACCCGCCGGGGAGCTCAAGAAGT 621

Qy    121 TCAAGCTGAAGCTGCTGTCGGGTCCGCGTGCGGAGGGGCTACGGGGCGCATCCCGCGGGGCG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    620 TCAAGCTGAAGCTGCTGTCGGGTCCGCGTGCGGAGGGGCTACGGGGCGCATCCCGCGGGGCG 561

Qy    181 CGCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 CGCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA 501

Qy    241 CCTACGCGCGCCAAGCTACCGTAACGTGCTGGGCGACATGGGCCTGCAGGGAGCTGCAGAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500 CCTACGCGCGCCAAGCTACCGTAACGTGCTGGGCGACATGGGCCTGCAGGGAGCTGCAGAGT 441

Qy    301 GGCGAGCTTCCAGGCGGGGGCACGCACCAGGGCCTCTCGGAGGATCCAGGCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    440 GGCGAGCTTCCAGGCGGGGGCACGCACCAGGGCCTCTCGGAGGATCCAGGCCC 381

Qy    361 CTCCTCAGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    380 CTCCTCAGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA 321

Qy    421 TCCCGAGGGGTCACAAACGTTGAGTGGTCGCTCGTGCTGAGGGAAGGTTCCTGACGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    320 TCCGGAGGGGTCACAAACGTTGAGTGGTCGCTCGTGCTGAGGGAAGGTTCCTGACGG 261

Qy    481 ATGAGCAGTACCAGGCAGTGCCGGCCGAGCCACCACCAACCCAGCGGGGAAGCTCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 ATGAGCAGTACCAGGCAGTGCCGGCCGAGCCACCACCAACCCAGCGGGGAAGCTCT 201

Qy    541 TCAGTTTCACACCAGCCTGGAACTGGACCTGCCTCAGGAGGAGCCCTAAGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 TCAGTTTCACACCAGCCTGGAACTGGACCTGCCTCAGGAGGAGCCCTAAGGG 141

Qy    601 AGTCCCAGTCCTCACCTGGTGGAGGAGACCTGGAGGCGGAGCTGTGAGGGCTCCTTCCCAGCAACAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 AGTCCCAGTCCTACCTGGTGGAGGAGACCTGGAGGCGGAGCTGTGAGGGCTTCCTTCCCAGCAACAC 81

Qy    661 TCCGGTCAGCCCCTTGGGCAATCCCACCAAAATCATCCTGAAATCTGAATCTGATCTTTTTTATACACAAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 TCCGGTCAGCCCCTTGGCAATCCCACCAAAATCATCCTGAATCTGATCTTTTTATACACAAT 21

Qy    721 ATACGAAAAGCCAGCTTGAA 740
          ||||||||||||||||||||
Db     20 ATACGAAAAGCCAGCTTGAA 1
```

RESULT 6
US-10-295-981-48
Sequence 48, Application US/10295981
Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/841,879B
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-10-295-981-48

Query Match   100.0%; Score 740; DB 15; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGGGTCCGGCTGCAGCGGCGGGTGAGCGGCCGCAGCGGCCGGGATCCTTGAGCCATGGGGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGGTCCGGCTGCAGCGGCGGGTGAGCGGCCGCAGCGGCCGGGATCCTTGAGCCATGGGGC 60

Qy     61 GGCGCGCGACCGCCATCCTCGGATGCGCTGGAGAACCTGACCCGCCGGGGAGCTCAAGAAGT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGCGCGCGACCGCCATCCTCGGATGCGCTGGAGAACCTGACCCGCCGGGGAGCTCAAGAAGT 120

Qy    121 TCAAGCTGAAGCTGCTGTCGGGTCCGCGTGCGGAGGGGCTACGGGGCGCATCCCGCGGGGCG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAGCTGAAGCTGCTGTCGGGTCCGCGTGCGGAGGGGCTACGGGGCGCATCCCGCGGGGCG 180

Qy    181 CGCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGCTGCTGTCCATGGACGCCCTTGGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA 240

Qy    241 CCTACGGCGCCGAGCTCACCGGTTAACGTGCTGGGCGACATGGGCCTGCAGGGAGATGGCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTACGGCGCCGAGCTCACCGGTTAACGTGCTGGGCGACATGGGCCTGCAGGGAGATGGCCG 300

Qy    301 GGCAGCTGCAGGCGGCGGCCACGCACCAGGGCCAGCTTCTGGAGGCCCAGGGCCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGCAGCTGCAGGCGGCGGCCACGCACCAGGGCCAGCTTCTGGAGGATCCAGGGCCC 360

Qy    361 CTCCTCAGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCCTCAGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCTTA 420
```

```
Qy    421  TCGCGGAGGGTCACAAAACGTTGAGTGGTCTCGTGGATGCTCGTGAGGAAGGTCCTGACGG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TCGCGGAGGGTCACAAAACGTTGAGTGGTCTCGTGGATGCTCGTGAGGAAGGTCCTGACGG  480

Qy    481  ATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAAACCCAAGCAAGATGCGGAAGCTCT  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  ATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAAACCCAAGCAAGATGCGGAAGCTCT  540

Qy    541  TCAGTTTCACACACCAGCCCTGAAACTGACCTGCAAGGACTTGCTCTCCTTCAGGCCCTCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TCAGTTTCACACACCAGCCCTGAAACTGACCTGCAAGGACTTGCTCTCCTTCAGGCCCTCA  600

Qy    601  AGTCCCCAGTCCTTACCTGGTGGAGGAGGACTGGAGGCTGGAGGCTTCCTTCCCAGCCAACAC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  AGTCCCCAGTCCTTACCTGGTGGAGGAGGACTGGAGGCTGGAGGCTTCCTTCCCAGCCAACAC  660

Qy    661  TCCGGTCAGCCCCTGGCAATCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTTATACACAAT  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TCCGGTCAGCCCCTGGCAATCCACCAAAATCATCCTGAATCTGAATCTGATCTTTTTTATACACAAT  720

Qy    721  ATACGAAAAAGCCAGCTTGAA  740
           |||||||||||||||||||||
Db    721  ATACGAAAAAGCCAGCTTGAA  740
```

Query Match　　99.1%;　Score 733;　DB 12;　Length 745;
Best Local Similarity　100.0%;　Pred. No. 9.7e-195;
Matches　733;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

```
Qy      8  GGCTGCAGCGCGGGGTGAGCGCGCGGCGCGCGGGGATCCTGGACCCATGGGCGCGCGCGCG   67
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2  GGCTGCAGCGCGGGGTGAGCGCGCGGCGCGCGGGGATCCTGGACCCATGGGCGCGCGCGCG   61

Qy     68  CGACGCCATCCTGGATGCGGCTGCGAGGAACCTGACGGAGCTCAAGAGAGTTCAAGCT  127
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62  CGACGCCATCCTGGATGCGGCTGCGAGGAACCTGACGGAGCTCAAGAGAGTTCAAGCT  121

Qy    128  GAAGCTGCTGTCGGTGCCGCTGCCGCGAGGGGCTAGGGGCGGGCGGCCATCCCGCGGGGGGGGCGGGGCTGCT  187
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128  GAAGCTGCTGTCGGTGCCGCTGCCGCGAGGGGCTAGGGGCGGGCGGCCATCCCGCGGGGGGGGCGGGGCTGCT  187
```

```
Db    122  GAAGCTGCTGTCGGTGCCGCTGCCGCGAGGGGCGGGGCTACGCGGCGGCGCATCCCGCGGGGGGGGCGGGGCTGCT  181

Qy    182  GTCCATGGACGCCCTTGGACCTCACCGGACCAAGCTCGGTCAGTTTCACTTCACTTGACGG  241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  GTCCATGGACGCCCTTGGACCTCACCGGACCAAGCTCGGTCAGTTTCACTTCACTTGACGG  247

Qy    242  CGCCGAGCTCACCGGTAACGTGCTGCGACAATGGGAGGAATGGCCGGGGGGCAGCT  301
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  CGCCGAGCTCACCGGTAACGTGCTGCGACAATGGGAGGAATGGCCGGGGGGCAGCT  307

Qy    302  GCAGGCGGGCGCCACGGCACCAGGGCTCTTGGGGGATCCAGGGCCCCTCCTCA  361
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GCAGGCGGGCGCCACGGCACCAGGGCTCTTGGGGGATCCAGGGCCCCTCCTCA  367

Qy    362  GTCGGGCCAGCCAAGCCAGGGCCTGCCACTTTATAGACAACACTCGGTG  421
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  GTCGGGCCAGCCAAGCCAGGGCCTGCCACTTTATAGACAACACTCGGTG  427

Qy    422  GGTCACAAAACGTTGAGTGGTCTCGTGGATGCTCGTGAGGAAGGTCCTGACGGATGAGCA  481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  GGTCACAAAACGTTGAGTGGTCTCGTGGATGCTCGTGAGGAAGGTCCTGACGGATGAGCA  487

Qy    482  GTACCAGGCAGTGCGGGGCCGAGCCCACCAAGCAAGATGCGGAAGCTCTTCAGTTT  541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488  GTACCAGGCAGTGCGGGGCCGAGCCCACCAAGCAAGATGCGGAAGCTCTTCAGTTT  547

Qy    542  CACACCAGCCCTGGAACTGGACTGGACTTGCTCCTCCCAGGGCCCTAAGGGAGTCCCA  601
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548  CACACCAGCCCTGGAACTGGACTGGACTTGCTCCTCCCAGGGCCCTAAGGGAGTCCCA  607

Qy    602  GTCCTTACCTGGTGGAGGACCTGGAGCTGGAGGCTTCCTTCCCAGCAACACTCCGGTC  661
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608  GTCCTTACCTGGTGGAGGACCTGGAGCTGGAGGCTTCCTTCCCAGCAACACTCCGGTC  667

Qy    662  AGCCCCTGGCAATCCACCAAAATCATCCTGAATCTGATCTTTTTTATACACAATATACGA  721
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662  AGCCCCTGGCAATCCACCAAAATCATCCTGAATCTGATCTTTTTTATACACAATATACGA  721

Qy    722  AAGCCAGCTTGAA  740
           |||||||||||||
Db    722  AAGCCAGCTTGAA  734
```

Query Match　　99.1%;　Score 733;　DB 12;　Length 779;
Best Local Similarity　100.0%;　Pred. No. 9.8e-195;

Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    7 CGGCTGCACGGGGGTGAGCGGCGGCAGCCCGCGCGGGGGGATCCTTGGAGCCATGGGGGCGGGGGGCGC       66
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   47 CGGCTGCACGGGGGTGAGCGGCGGCAGCCCGCGCGGGGGGATCCTTGGAGCCATGGGGGCGGGGGGCGC      106

Qy   67 GCGACGCCCATCCTGATGCGCTGAGAACCTGACGCGCGGGGATCCAAGAGAGTTCAAGC                  126
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  107 GCGACGCCCATCCTGATGCGCTGAGAACCTGACGCGCGGGGATCCAAGAGAGTTCAAGC                 166

Qy  127 TGAAGCTGCTGTCGGTGCCCCTGCCCGAGGGGCTA...CGGGGGCATCCCGCGGGGGGGCTGC           186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  167 TGAAGCTGCTGTCGGTGCCCCTGCCCGAGGGGCTA...CGGGGGCATCCCGCGGGGGGGCTGC          226

Qy  187 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTTCTTACCTGGGGACCTACG                      246
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  227 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTTCTTACCTGGGGACCTACG                     286

Qy  247 GCGCCGAGCTCACCGGCTAACGTGCTGCGCGACATGGGCGTCTGCAGGGCCAGC                        306
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  287 GCGCCGAGCTCACCGGCTAACGTGCTGCGCGACATGGGCGTCTGCAGGGCCAGC                       346

Qy  307 TGCAGGCGGCGCACCGGGCTCTTGGAGCGCGCGCCGAGCCCAGCCCCTCCTC                          366
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  347 TGCAGGCGGCGCACCGGGCTCTTGGAGCGCGCGCCGAGCCCAGCCCCTCCTC                         406

Qy  367 AGTCGGCACCCAACCCAGCCTGCACTTTATAGACCAGCAGCACCTTATCGCGA                        426
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  407 AGTCGGCACCCAACCCAGCCTGCACTTTATAGACCAGCAGCACCTTATCGCGA                       466

Qy  427 GGGTCACAAAACGTTGAGTGGCGCTGCTCTTGTACGGGATGCTCTTGTACGGGATGAGC                  486
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  467 GGGTCACAAAACGTTGAGTGGCGCTGCTCTTGTACGGGATGCTCTTGTACGGGATGAGC                 526

Qy  487 AGTACCAGGCGAGTGCGGGGCCGAGCCCACCCAAGCAAGATGGGGAAGCTCTTCAGTT                   546
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  527 AGTACCAGGCGAGTGCGGGGCCGAGCCCACCCAAGCAAGATGGGGAAGCTCTTCAGTT                  586

Qy  547 TCACACCAGCCTGGAACTGACCTTGCTCCTCCAGGCGCCCTTAAGGGAGTCCC                        606
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  587 TCACACCAGCCTGGAACTGACCTTGCTCCTCCAGGCGCCCTTAAGGGAGTCCC                       646

Qy  607 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGGTCTTCCTTCCCAGGCAACACTCCGGT            666
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  647 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGGTCTTCCTTCCCAGGCAACACTCCGGT           706

Qy  667 CAGCCCCTGGCAATCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA               726
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  707 CAGCCCCTGGCAATCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA              766

Qy  727 AAAGCCAGCTTGA                                                                739
        |||||||||||||
Db  767 AAAGCCAGCTTGA                                                               779
```

Query Match 97.5%; Score 721.2; DB 15; Length 806;
Best Local Similarity 99.6%; Pred. No. 2e-191;
Matches 731; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

```
Qy    7 CGGCTGCACGGGGGTGAGCGGCGGCAGCCCGCGCGGGGGGATCCTTGGAGCCATGGGGGCGGGGGGCGC       66
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  739 CGGCTGCACGGGGGTGAGCGGCGGCAGCCCGCGCGGGGGGATCCTTGGAGCCATGGGGGCGGGGGGCGC      680

Qy   67 GCGACGCCCATCCTGATGCGCTGAGAACCTGACGCGCGGGGATCCAAGAGAGTTCAAGC                  126
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  679 GCGACGCCCATCCTGATGCGCTGAGAACCTGACGCGCGGGGATCCAAGAGAGTTCAAGC                 620

Qy  127 TGAAGCTGCTGTCGGTGCCCCTGCCCGAGGGGCTACGGGGGCATCCCGCGGGGGGGCTGC                 186
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  619 TGAAGCTGCTGTCGGTGCCCCTGCCCGAGGGGCTACGGGGGCATCCCGCGGGGGGGCTGC                560

Qy  187 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTTCTTACCTGGGGACCTACG                      246
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  559 TGTCCATGGACGCGCCTTGGACCTCACCGACAAGCTTCTTACCTGGGGACCTACG                     500

Qy  247 GCGCCGAGCTCACCGGCTAACGTGCTGCGCGACATGGGCGTCTGCAGGGCCAGC                        306
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  499 GCGCCGAGCTCACCGGCTAACGTGCTGCGCGACATGGGCGTCTGCAGGGCCAGC                       440

Qy  307 TGCAGGCGGCGCACCGGGCTCTTGGAGCGCGCGCCGAGCCCAGCCCCTCCTC                          366
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  439 TGCAGGCGGCGCACCGGGCTCTTGGAGCGCGCGCCGAGCCCAGCCCCTCCTC                         380

Qy  367 AGTCGGCACCCAACCCAGCCTGCACTTTATAGACCAGCAGCACCTTATCGCGA                        426
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 AGTCGGCACCCAACCCAGCCTGCACTTTATAGACCAGCAGCACCTTATCGCGA                       320

Qy  427 GGGTCACAAAACGTTGAGTGGCGCTGCTCTTGTACGGGATGAGC                                 486
        ||||||||||||||||||||||||||||||||||||||||||||
Db  319 GGGTCACAAAACGTTGAGTGGCGCTGCTCTTGTACGGGATGAGC                                260

Qy  487 AGTACCAGGCGAGTGCGGGGCCGAGCCCACCCAAGCAAGATGGGGAAGCTCTTCAGTT                   546
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 AGTACCAGGCGAGTGCGGGGCCGAGCCCACCCAAGCAAGATGGGGAAGCTCTTCAGTT                  200

Qy  547 TCACACCAGCCTGGAACTGACCTTGCTCCTCCAGGCGCCCTTAAGGGAGTCCC                        606
        ||||||||||||||||||||||||||||||||||||::||||||||||||||
Db  199 TCACACCAGCCTGGAACTGACCTTGCTCCTCCAGGCCCCTTAAGGGAGTCCC                        140

Qy  607 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGGTCTTCCTTCCCAGGCAACACTCCGGT            666
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  139 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGGTCTTCCTTCCCAGGCAACACTCCGGT            80

Qy  667 CAGCCCCTGGCAATCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA               726
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 CA-CCCCTGGCAATCCCACCCAAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA               21

Qy  727 AAAGCCAGCTTGA                                                                739
        |||||||||||||
Db   20 AAAGCCAGCTTGA                                                                 7
```

Left column:

```
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 278
LENGTH: 811
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-301-278
```

Query Match   94.3%;   Score 697.6; DB 9; Length 811;
Best Local Similarity   99.5%; Pred. No. 7.6e-185;
Matches 730; Conservative 1; Mismatches 0; Indels 3; Gaps 3;

```
Qy      7 CGGCTGCAGCGGGGGTGAGGGGCCGGCAGCGGGGGGATCCTGAGCCATGGGGGGGGGCGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 CGGCTGCAGCGGGGGTGAGGGGCAGCGGGGGGATCCTGAGCCATGGGGGGGGGCGC 127

Qy      7 GGCAACGGCCATCCTTGGGATCGGCTGGAGAACCTGACCGCGGCTCAAGAAGTTCAAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 GGCAACGGCCATCCTTGGGATCGGCTGGAGAACCTGACCGCGGCTCAAGAAGTTCAAGC 126

Qy    128 GCCGACGCCATCCTTGGGATGGGCTGGAGAACCTGACCGCGGCTCAAGAAGTTCAAGC 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 TGAAGCTTGCGTGCGGTGCGGCTACGGCGCCATCCCGCGGGGGGGGCGGCTGC 186

Qy    188 TGAAGCTGCGTGCGGTGCGGCTACGGCGCCATCCCGCGGGGGGGGCGGCTGC 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TGTCCATGGACGGCCTTGGAACTCACCGACAAGTCGGTCAGCTTCTACCTGGACGACCTACG 307

Qy    247 GCGCCGAGCTCACCGCTACGTCGTGCCGCGACCATGGGGCCTGCAGGGAGGATTGGCGCGGCAGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 GCGCCGAGCTCACCGCTACGTCGTGCCGCGACCATGGGGCCTGCAGGGAGGATTGGCGCGGCAGC 367

Qy    307 TGCAGGGCGGCCACGGCACCAGGGCGGGGTCTGGAGCCCGCGGCCGCCCCCTCCTC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 TGCAGGGCGGCCACGGCACCAGGGCGGGGTCTGGAGCCCGCGGCCGCCCCCTCCTC 427

Qy    367 AGTCGGCAGCCAGCCAGGCCCTGCACTTTATAGACCAGCAGCTGCGCTTATCGCGCA 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 AGTCGGCAGCCAAGCCAGGCCCTGCACTTTATAGACCAGCAGCTGCGCTTATCGCGCA 487

Qy    427 GGGGTCACAAACGTTGAGTGGGCTCCTGGATGGCTCCTGACGGATGGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 GGGGTCACAAACGTTGAGTGGGCTCCTGGGAAGGTCCTGACGGATGGAGC 547

Qy    487 AGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAAGATGGGGAAGCTGCTTCTTCAGTT 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    548 AGTACCAGGCAGTGC-GGCCGAGCCCACCAACCCAAGCAAGGAAGCTGCTTCTTCAGTT 606

Qy    607 AGTCCTACCTGGTGGAGAGCGAGTCTGAGGCGTCCTCGGGCTCCTCCCAGCGAAACACTCCCGT 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    667 AGTCCTACCTGGTGGAGAGACCTGAGGCGTCCTCGGGCTCCTCCCAGCGAAACACTCCCGT 725

Qy    667 CAGCCCCTGGCAATCCACCAACAATCACTCCTGAATCTGATCTTTTTATACACAATATACGA 726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 CA-SCCTGGCAATCCCACCAAATCATCCTGAATCTGATCTTTTTATACACAATATACGA 784

Qy    727 AAAGCCAGCTTGAA 740
          ||||||||||||||
Db    785 AAAGCCAGCTTGAA 798
```

Right column:

```
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREO
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/728,721
CURRENT FILING DATE:
PRIOR APPLICATION NUMBER: 2000-12-01
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/340,620
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 50
LENGTH: 585
TYPE: DNA
ORGANISM: Homo sapiens
US-09-728-721-50
```

Query Match   79.1%;   Score 585; DB 9; Length 585;
Best Local Similarity   100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     54 ATGGGGCGCGGGCGCCGGACCCGCCATCCTTGGGATGCGCTGGAGAAACCTGACCGCGGCTC 113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGGCGCGGGCGCCGGACCCGCCATCCTTGGGATGCGCTGGAGAAACCTGACCGCGGCTC 60

Qy    114 AAGAAGTTCAAGCTGAAGCTGCGTGTCGGTGCGGTGCGGCTACGGCGCCATCCCG 173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAGAAGTTCAAGCTGAAGCTGCGTGTCGGTGCGGTGCGGCTACGGCGCCATCCCG 120

Qy    174 CGGGCGCGCGCTGCTGTCCATGGACGGCCTTGGAACTCACCGACAAGCTGGTCAGCTTCTAC 233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGGGCGCGCGCTGCTGTCCATGGACGGCCTTGGAACTCACCGACAAGCTGGTCAGCTTCTAC 180

Qy    234 CTGGAGACCTACGCGCGCGCCAGCGTCACCCGCTAAACGTGCTCGCGCGACACATGGGCGTGCAGGAG 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGGAGACCTACGCGCGCGCCAGCGTCACCCGCTAAACGTGCTCGCGCGACACATGGGCGTGCAGGAG 240

Qy    294 ATGGCCGGGCAGCTGCAGGGCGGCCACGGCACCAGGGCGGGGTCTGGAGCCGCTGGGATC 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATGGCCGGGCAGCTGCAGGGCGGCCACGGCACCAGGGCGGGGTCTGGAGCCGCTGGGATC 300

Qy    354 CAGGCCCCTCCTCAGTCGGCGGCAAGCGCAAGGCCAGGCCTGCACTTTATAGACCAGCAGCCACGGCT 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGGCCCCTCCTCAGTCGGCGGCAAGCGCAAGGCCAGGCCTGCACTTTATAGACCAGCAGCCACGGGCT 360

Qy    414 GCGCTTATCGCGAGGGTCACAAAACGTTGAGTGGTTGCTGGATGCTCTGTTACGGGGAAGGTC 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCGCTTATCGCGAGGGTCACAAAACGTTGAGTGGTTGCTGGATGCTCTGTTACGGGGAAGGTC 420

Qy    474 CTGACGGATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGAGATGCGG 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGACGGATGAGCAGTACCAGGCAGTGCGGGGCCGAGCCCACCAACCCAAGCAAGAGATGCGG 480

Qy    534 AAGCTCTTCAGTTTCACACCAGCAGCTGGGAACTGGACCTGCAAGGAGGCC 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAGCTCTTCAGTTTCACACCAGCAGCTGGGAACTGGACCTGCAAGGAGGCC 540

Qy    594 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGCGGGAGC 638
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGCGGGAGC 585
```

TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-340001
CURRENT APPLICATION NUMBER: US/09/996,617
CURRENT FILING DATE: 2001-11-27
PRIOR APPLICATION NUMBER: 09/931,071
PRIOR FILING DATE: 2001-10-27
PRIOR APPLICATION NUMBER: 09/428,252
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-10-27
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 9
  LENGTH: 585
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-996-617-9

Query Match   79.1%;   Score 585;   DB 10;   Length 585;
Best Local Similarity 100.0%;   Pred. No. 2e-153;
Matches 585;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy   54 ATGGGGGCGCGCGCCGGCGACCCATCCTGGATGCGCTGGAGAAACCTGACCCGCCGAGGAGCTC 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGGGGCGCGCGCCGGCGACCCATCCTGGATGCGCTGGAGAAACCTGACCCGCCGAGGAGCTC  60

Qy  114 AAGAAGTTCAAGCTGAAGCTGCTGTCCGGTGCCGCTGCGGGAGGGGCTACCGGCGCCATCCCG 173
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAGAAGTTCAAGCTGAAGCTGCTGTCCGGTGCCGCTGCGGGAGGGGCTACCGGCGCCATCCCG 120

Qy  174 CGGGGGCGCGCTGCTTGCTTGGACGCCCTTGGACGCCAAGACTCGGTGCAGCTTCTAC 233
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CGGGGGCGCGCTGCTTGCTTGGACGCCCTTGGACGCCAAGACTCGGTGCAGCTTCTAC 180

Qy  234 CTGGAGAACCTTACGGCGCCGCCCAGCTCACCGGTTAACCGTGCTGCTGCGCGACATGGGCGAG 293
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGGAGAACCTTACGGCGCCGCCCAGCTCACCGGTTAACCGTGCTGCTGCGCGACATGGGCGAG 240

Qy  294 ATGGCCGGCGCAGCTGCAGCGCGGCCGCCACGCGACCACCAGGGCTCTCGGAGCGGGCGGGATC 353
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ATGGCCGGCGCAGCTGCAGCGCGGCCGCCACGCGACCACCAGGGCTCTCGGAGCGGGCGGGATC 300

Qy  354 CAGGCCCCTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCGGGTCTCGTGTACGGGAAGGTC 413
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGGCCCCTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCGGGTCTCGTGTACGGGAAGGTC 360

Qy  414 GCGCTTATCGGCGAGGGGTCACAAAACGTTGAGTGGGCTGCTGCGATGCTCTGTTACGGGAAGGTC 473
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCGCTTATCGGCGAGGGGTCACAAAACGTTGAGTGGGCTGCTGCGATGCTCTGTTACGGGAAGGTC 420

Qy  474 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCCGCCACCAACCCAAGCAAGATGCGG 533
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCCGCCACCAACCCAAGCAAGATGCGG 480

Qy  534 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCC 593
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCC 540

Qy  594 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGACGGGAGC 638
        |||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGACGGGAGC 585
```

RESULT 13
US-09-841-879B-6
Sequence 6, Application US/09841879B
Patent No. US20020142979A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-333001
CURRENT APPLICATION NUMBER: US/09/841,879B

CURRENT FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: US 09/728,721
PRIOR FILING DATE: 2000-12-01
CURRENT APPLICATION NUMBER: US 09/340,620
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
  LENGTH: 585
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-841-879B-6

Query Match   79.1%;   Score 585;   DB 10;   Length 585;
Best Local Similarity 100.0%;   Pred. No. 2e-153;
Matches 585;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy   54 ATGGGGGCGCGCGCCGGCGACCCATCCTGGATGCGCTGGAGAAACCTGACCCGCCGAGGAGCTC 113
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGGGGCGCGCGCCGGCGACCCATCCTGGATGCGCTGGAGAAACCTGACCCGCCGAGGAGCTC  60

Qy  114 AAGAAGTTCAAGCTGAAGCTGCTGTCCGGTGCCGCTGCGGGAGGGGCTACCGGCGCCATCCCG 173
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAGAAGTTCAAGCTGAAGCTGCTGTCCGGTGCCGCTGCGGGAGGGGCTACCGGCGCCATCCCG 120

Qy  174 CGGGGGCGCGCTGCTTGCTTGGACGCCCTTGGACGCCAAGACTCGGTGCAGCTTCTAC 233
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CGGGGGCGCGCTGCTTGCTTGGACGCCCTTGGACGCCAAGACTCGGTGCAGCTTCTAC 180

Qy  234 CTGGAGAACCTTACGGCGCCGCCCAGCTCACCGGTTAACCGTGCTGCTGCGCGACATGGGCGAG 293
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGGAGAACCTTACGGCGCCGCCCAGCTCACCGGTTAACCGTGCTGCTGCGCGACATGGGCGAG 240

Qy  294 ATGGCCGGCGCAGCTGCAGCGCGGCCGCCACGCGACCACCAGGGCTCTCGGAGCGGGCGGGATC 353
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ATGGCCGGCGCAGCTGCAGCGCGGCCGCCACGCGACCACCAGGGCTCTCGGAGCGGGCGGGATC 300

Qy  354 CAGGCCCCTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCGGGTCTCGTGTACGGGAAGGTC 413
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGGCCCCTCCTCAGTCGGCGCAGCCAAGCCAGGCCTGCGGGTCTCGTGTACGGGAAGGTC 360

Qy  414 GCGCTTATCGGCGAGGGGTCACAAAACGTTGAGTGGGCTGCTGCGATGCTCTGTTACGGGAAGGTC 473
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCGCTTATCGGCGAGGGGTCACAAAACGTTGAGTGGGCTGCTGCGATGCTCTGTTACGGGAAGGTC 420

Qy  474 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCCGCCACCAACCCAAGCAAGATGCGG 533
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGACGGATGAGCAGTACCAGGCAGTGCGGGCCGCCACCAACCCAAGCAAGATGCGG 480

Qy  534 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCC 593
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AAGCTCTTCAGTTTCACACCAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCC 540

Qy  594 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGACGGGAGC 638
        |||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGACGGGAGC 585
```

RESULT 14
US-10-295-981-50
Sequence 50, Application US/10295981
Publication No. US20030120055A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/10/295,981
CURRENT FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US/09/340,620
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359

```
;    PRIOR FILING DATE: 1998-12-08
;    PRIOR APPLICATION NUMBER: US 09/099,041
;    PRIOR FILING DATE: 1998-06-17
;    PRIOR APPLICATION NUMBER: US 09/019,942
;    PRIOR FILING DATE: 1998-02-06
;    NUMBER OF SEQ ID NOS: 71
;    SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 50
;    LENGTH: 585
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-295-981-50

Query Match          79.1%; Score 585; DB 15; Length 585;
Best Local Similarity   100.0%; Pred. No. 2e-153;
Matches  585; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy     54 ATGGGGCGGGCGCTGTCGGGGATCGCGTCGGAGGAACCTGACCGCGAGGAGCTC 113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGGCGGGCGCTGTCGGGGATCGCGTCGGAGGAACCTGACCGCGAGGAGCTC 60

Qy    114 AAGAAGTTCAAGCTGAAGCTGCTGTGTCGGGTGCCGGCGCTACGGGCGCATCCCG 173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAGAAGTTCAAGCTGAAGCTGCTGTGTCGGGTGCCGGCGCTACGGGCGCATCCCG 120

Qy    174 CGGGGGCGGGCGCTGTCCATGGACGGCCTTGGACCTCACAGAGCTGGTCAGCTTCTAC 233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGGGGGCGGGCGCTGTCCATGGACGGCCTTGGACCTCACCGAGCTGGTCAGCTTCTAC 180

Qy    234 CTGGAGACCTACGGGCCCGAGCTCACCGCGCCGAGCTCAAGGTGCTTACGGGCCTGCAGGAG 293
          |||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGGAGACCTACGGGCCCGAGCTCACCGCGCCGAGCTCAAGGTGCTTACGGGCCTGCAGGAG 240

Qy    294 ATGGGCCGGGCAGCTGCAGGGGCCACGCACCAGGGGCTCTCGGAGCCCGGCGACCTGGGATC 353
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATGGGCCGGGCAGCTGCAGGGGCCACGCACCAGGGGCTCTCGGAGCCCGGCGACCTGGGATC 300

Qy    354 CAGGGCCCCTCCTCAGTCGGCAGCCAAGCCAGGCCTGCACTTTATAGACAGCACACCGGGGCT 413
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGGGCCCCTCCTCAGTCGGCAGCCAAGCCAGGCCTGCACTTTATAGACAGCACACCGGGGCT 360

Qy    414 GCGCTTATCGGCGAGGGGTCACAAACGTTGAGTGGCTGCTGTACGGGAAGGTC 473
          ||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCGCTTATCGGCGAGGGGTCACAAACGTTGAGTGGCTGCTGTACGGGAAGGTC 420

Qy    474 CTGACGGATGAGCAGTACACCAGGCCAGTGCGGGCCGGAGCCCAACCCAAGCAAGATGCGG 533
          |||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGACGGATGAGCAGTACACCAGGCCAGTGCGGGCCGGAGCCCAACCCAAGCAAGATGCGG 480

Qy    534 AAGCTCTTCAGTTTCACACCAGCCTGGACTTGCTCCTCCAGGCC 593
          ||||||||||||||||||||||||||||||||||||||||
Db    481 AAGCTCTTCAGTTTCACACCAGCCTGGACTTGCTCCTCCAGGCC 540

Qy    594 CTAAGGGAGTCCCAGTCCTACCTGGTTGGAGGACCTGGAGCGGAGC 638
          ||||||||||||||||||||||||||||||||||||||||||
Db    541 CTAAGGGAGTCCCAGTCCTACCTGGTTGGAGGACCTGGAGCGGAGC 585
```

```
;    PRIOR FILING DATE: 1999-11-03
;    NUMBER OF SEQ ID NOS: 8564
;    SOFTWARE: Patentin Ver. 3.0
;  SEQ ID NO 1145
;    LENGTH: 639
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-106-698-1145

Query Match          71.8%; Score 531; DB 15; Length 639;
Best Local Similarity   94.9%; Pred. No. 2.4e-138;
Matches  595; Conservative  0; Mismatches  20; Indels  12; Gaps  4;

Qy    124 AGCTGAAGCTGCTGTCGGTCCGGTGCCGGTGCCGGCGGGGCTACGGGCGCCATCCCGCG------G 176
          |||||||||||||||||||||||||||||||||||||||||||
Db      2 AGCTGAAAGCTGCTGCTGTGTCGGTCGGGTGCCGGGCGGGGCTACGGGCGGCGGACGCCATCCTG 61

Qy    177 GGGCGCGGCTGCTGTCCATGGACGGCCCTTGGACCTCACCGA---CAAGCTGGTTCAGCTTCTAC 233
          |||||||||||||||||||||||||||||||||||
Db     62 GATGCGCGCTGGAGAACCTGACCGCCGAGGAGCTTCAAGAAGTTCAAGCTGGTGCTTCGTAC 121

Qy    234 CTGGAGACCTACGGGCCCGGAGCTCACCGCGCCGAGCTCAAGGTGCTGCTTGCGGGCCTGCAGGAG 293
          |||||||||||||||||||||||||||||||||||||||||||
Db    122 CTGGAGACCTACGGGCGCGGAGCTTCACGGCTTAACGTTGCTGCGGCGGACAATGGGCCTGCAGGAG 181

Qy    294 ATGGGCCGGGCAGCTGCAGGGGCGACCCACCGCACCCAGGGGCTCTCGGAGCCCGGCCGACCTGGGATC 353
          ||||||||||||||||||||||||||||||||||||||
Db    182 ATGGGCCGGGCAGCTGCAGGGCGCCACCACCAGGGGCTCTCGGAGCCCGGCCGACCTGGGATC 241

Qy    354 CAGGGCCCCTCCTCAGTCGGCCAGGCAGGCCGAGGCCAGTGC-GGCGAGGGCCAACCCAAGCAAGGATCGGG 413
          ||||||||||||||||||||||||||||||||||||||||||
Db    242 CAGGGCCCCTCCTCAGTCGGCAGCAAACCAGGCCAGTGCC-GGCGAGGCCGAGCCCAACCCAAGCAAGGATGCGG 301

Qy    414 GCGCTTATCGGCCGAGGGGTCACAAAACGTTGAGTGGGTGCTGCTGTACGGGGAAGGTC 473
          ||||||||||||||||||||||||||||||||||||||||
Db    302 GCGCTTATCGGCCGAGGGGTCACAAAACGTTGAGTGGCTGCTGTACGGGGAAGGTC 361

Qy    474 CTGACGGATGAGCAGTACCAGCAGTCCAGCGGGGCCGGGGAGCCAACCCAAGCAAGGATGGGG 533
          |||||||||||||||||||||||||||||||||||||||||||
Db    362 CTGACGGATGAGCAGTACCAGCAGTCCAGGCCAGGTCC--GGCGAGGCCGAGCCCAACCCAAGCAAAGATGCGG 420

Qy    534 AAGCTCTTCAGTTTCACACCAGCCTGGACTTGCTCAAGGACACTGGACTTGCTCCTCCAGGCC 593
          |||||||||||||||||||||||||||||||||||||||||||
Db    421 AAGCTCTTCAGTTTCACACCAGCCTGGACTTGCTCCTCCAGGCC 480

Qy    594 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGCGGAGCTTCCTTCCCA 653
          |||||||||||||||||||||||||||||||||||||||||||
Db    481 CTAAGGGAGTCCCAGTCCTACCTGGTGGAGGACCTGGAGCGGAGCTTCCTTCCCA 540

Qy    654 GCAACACTCCGGTCACGCCCTGGCAATCCCACCAAAATCATCCTGAATCTTTTTTAT 713
          |||||||||||||||||||||||||||||||||||||||||
Db    541 GCAACACTCCGGTCAG-CCCTGGCCAATCCCACCAAAATCATCCTGAATCTTTTTTAT 599

Qy    714 ACACAATATACGAAAAAGCCCAGCTTGAA 740
          |||||||||||||||||||||||||||
Db    600 ACACAATATACGAAAAAGCCCAGCTTGAA 626
```

```
RESULT 15
US-10-106-698-1145
;  Sequence 1145, Application US/10106698
;  Publication No. US20030109690A1
;GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;  FILE REFERENCE: PA005P1
;  CURRENT APPLICATION NUMBER: US/10/106,698
;  CURRENT FILING DATE: 2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
```

Search completed: January 29, 2004, 13:44:37
Job time : 236.019 secs

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 9.48584 seconds
                                         (without alignments)
                                         3262.936 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTAEEL............LLLQALRESQSYLVEDLERS 195

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : A_Geneseq 19Jun03:*
   1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
   2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
   3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
   4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
   5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
   6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
   7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
   8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
   9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
  10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
  11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
  12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
  13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
  14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
  15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
  16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
  17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
  18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
  19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
  20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
  21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
  22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
  23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
  24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	990	100.0	195	20	AAY48553	Human breast tumou
2	990	100.0	195	22	AAU58525	Human novel cytoki
3	990	100.0	195	22	AAB00588	Human target of me
4	990	100.0	195	22	AAB20085	Human CARD-5 prote
5	990	100.0	195	23	AAU99353	Human caspase recr
6	990	100.0	195	23	AAO17854	Pyrin domain conta
7	990	100.0	195	24	ABG71635	Human caspase recr
8	990	100.0	205	22	AAG74647	Human colon cancer
9	873.5	88.2	176	22	AAB00594	Alternatively spli

No.	Score	%	Length		DB ID	Description
10	759	76.7	190	21	AAB43675	Human cancer assoc
11	684	69.1	193	22	AAB00592	Mouse target of me
12	684	69.1	193	22	AAB20086	Mouse CARD-5 prote
13	684	69.1	193	23	AAU99352	Mouse caspase recr
14	576	58.2	171	22	AAE00593	Rat target of meth
15	564.5	57.0	158	22	AAG74648	Human colon cancer
16	545	55.1	110	22	AAE00595	Alternatively spli
17	509	51.4	136	21	AAB53881	Human colon cancer
18	448	45.3	91	22	AAE00589	Human target of me
19	439	44.3	84	22	AAE00591	Human target of me
20	305	30.8	76	23	AAO21934	PAAD domain-contai
21	267	27.0	89	22	AAB36608	Human FLXHT-30 pr
22	267	27.0	89	23	AAO17852	Pyrin domain conta
23	253	25.6	1397	22	AAY72670	Human NB-ARC and C
24	250.5	25.3	1429	22	AAB62571	Human CARD-7 polyp
25	250.5	25.3	1429	23	ABG97969	Human NB-ARC and C
26	250.5	25.3	1429	23	ABG78472	Human leucine rich
27	250.5	25.3	1429	23	ABB77916	Human leucine-rich
28	250.5	25.3	1429	24	ABG71631	Human leucine-rich
29	250.5	25.3	1429	24	ABG71633	Human caspase recr
30	250	25.3	1442	22	AAY72671	Human caspase recr
31	250	25.3	1473	22	AAY72669	Human NB-ARC and C
32	250	25.3	1473	22	AAY72711	Human NAC beta iso
33	249.5	25.2	1429	22	ABG78455	Human caspase recr
34	249.5	25.2	1429	23	AAO17855	Pyrin domain conta
35	249.5	25.1	1473	22	AAO17856	Human G-protein co
36	248.5	24.2	442	22	AAB24513	Human G-protein co
37	240	18.4	76	22	AAE06758	Human G-protein co
38	182	15.5	65	21	AAO21935	PAAD domain-contai
39	153	14.8	77	22	AAB24519	Human secreted pro
40	137	13.8	77	22	AAE07516	Human PYRIN-1 doma
41	125.5	12.7	891	22	AAE07515	Human PYRIN-1 doma
42	125.5	12.7	1851	22	AAU00023	Human activated T-
43	125.5	12.7	1851	23	AAO17860	Pyrin domain conta
44	122.5	12.4	781	20	AAY09001	FMF associated pro
45	122.5	12.4	781	23	AAO17853	Pyrin domain conta

ALIGNMENTS

RESULT 1
AAY48553

ID   AAY48553 standard; Protein; 195 AA.
XX
AC   AAY48553;
XX
DT   08-DEC-1999 (first entry)
XX
DE   Human breast tumour-associated protein 14.
XX
KW   Expressed sequence tag; EST; human; breast; cancer; gene therapy;
KW   treatment; tumour; cytostatic; medicament.
XX
OS   Homo sapiens.
XX
PN   DE19813839-A1.
XX
PD   23-SEP-1999.
XX
PF   20-MAR-1998; 98DE-1013839.
XX
PR   20-MAR-1998; 98DE-1013839.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;
XX
DR   WPI; 1999-528981/45.
DR   N-PSDB; AAZ33631.
XX
PT   Human nucleic acid sequences and protein products from tumor breast
PT   tissue, useful for breast cancer protein therapy -

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
XX
PS
CC   This invention describes novel human nucleic acid sequences from tumor
CC   breast tissue which have cytostatic activity. The nucleic acid sequences
CC   can be used to produce and isolate full-length gene sequences. They can
CC   be used to express proteins, which can be used as tools to find an
CC   activity against breast cancer. The sequences can be used in sense or
CC   antisense form. They are especially useful for medicaments for gene
CC   therapy to treat breast cancer. AaY48540-Y48617 represent protein
CC   fragments encoded by the expressed sequence tags described in the method
CC   of the invention.
XX
SQ   Sequence   195 AA;

Query Match             100.0%;  Score 990; DB 20; Length 195;
Best Local Similarity   100.0%;
Matches 195; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy      1 MGRARDAILDALENLTAEELKKFFLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRARDAILDALENLTAEELKKFFLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  60

Qy     61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy    121 ALIARVTNVEWLLDALYGKVLTDDEQYQAVRAEPTNPSKQWRKLFSFTPAWNWTCKDLLLQA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy    181 LRESQSYLVEDLERS 195
          ||||||||||||||
Db    181 LRESQSYLVEDLERS 195

RESULT 2
ID   AAU68525
AC   AAU68525;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human novel cytokine encoded by cDNA 79OCIP2B_1 #1.
XX
KW   Human; cytokine; cell proliferation; cell differentiation;
KW   antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
KW   nervous system disease; neuropathy; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; spinal cord disorder;
KW   head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
KW   platelet disorder; thrombocytopaenia; stem cell disorder;
KW   aplastic anaemia; tissue regeneration; wound healing; ulcer;
KW   osteoporosis; osteoarthritis; bone degenerative disorder;
KW   periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
KW   severe combined immunodeficiency; infection; autoimmune disorder;
KW   multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
KW   asthma; coagulation disorder; haemophilia; sepsis; nephritis;
KW   inflammatory bowel disease; food supplement; immunogen.
XX
OS   Homo sapiens.
XX
PN   WO200175093-A1.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US10484.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
PR   22-SEP-2000; 2000US-0668680.
PR   23-OCT-2000; 2000US-0695618.
PR   30-NOV-2000; 2000US-0728711.
```

```
PR   14-MAR-2001; 2000US-0728711.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;
PI   Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;
XX
DR   WPI; 2001-626432/72.
DR   N-PSDB; AAS59817.
XX
PT   New polypeptides and nucleic acids, useful for diagnosis, treatment of
PT   inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
PT   degenerative disorders, cancer and promoting wound healing -
XX
PS   Claim 20; Page 242-243; 336pp; English.
XX
CC   The invention relates to isolated human polypeptides (which may be
CC   cytokines) and the polynucleotides encoding them. The protein is useful
CC   for identifying a compound which binds to it (e.g. modulators, agonists
CC   and antagonists). The polynucleotides are useful as an array for mismatch
CC   detection. The proteins and nucleic acids are useful as nutritional
CC   sources or supplements. The protein exhibits activity relating
CC   to cytokine, cell proliferation, cell differentiation, antiinflammatory,
CC   stem cell growth factor activity, activin stimulating or immune
CC   suppressing and activin or inhibin related activities. The proteins (and
CC   antibodies raised against them) and nucleic acids are therefore useful in
CC   the diagnosis and treatment of diseases and disorders such as cancer,
CC   central and peripheral nervous system diseases and neuropathies,
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
CC   diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
CC   thrombocytopaenia, stem cell disorders, aplastic anaemia, for
CC   regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC   growth, and in tissue repair, healing of burns, incisions, ulcers, for
CC   treating osteoporosis, osteoarthritis, bone degenerative disorders, or
CC   periodontal disease, lung or liver fibrosis, reperfusion injury in
CC   various tissues, various immune deficiencies and disorders including
CC   severe combined immunodeficiency (SCID), bacterial or fungal infections,
CC   autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
CC   diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
CC   such as asthma or other respiratory problems, coagulation disorders,
CC   haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory
CC   bowel disease, viral infection and are useful in altering bodily
CC   characteristics. The present sequence represents a novel protein of the
CC   invention.
XX
SQ   Sequence   195 AA;

Query Match             100.0%;  Score 990; DB 22; Length 195;
Best Local Similarity   100.0%;
Matches 195; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy      1 MGRARDAILDALENLTAEELKKFFLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRARDAILDALENLTAEELKKFFLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  60

Qy     61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy    121 ALIARVTNVEWLLDALYGKVLTDDEQYQAVRAEPTNPSKQWRKLFSFTPAWNWTCKDLLLQA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy    181 LRESQSYLVEDLERS 195
          ||||||||||||||
Db    181 LRESQSYLVEDLERS 195

RESULT 3
ID   AAE00588
AC   AAE00588;
XX
AAE00588 standard; Protein; 195 AA.
```

AC   AAE00588;
XX
DT   02-JUL-2001 (first entry)
XX
DE   Human target of methylation-induced silencing-1 (TMS1) protein.
XX
KW   Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast.
XX
OS   Homo sapiens.
XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999; 99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   vertino PM;
XX
PS   Claim 85; Page 114; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it. The present
CC   sequence is human target of methylation-induced silencing-1
CC   (TMS1) protein.
XX
SQ   Sequence 195 AA;

DR   WPI; 2001-290922/30.
DR   N-PSDB; AAD03889, AAD03890.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation useful for identifying subject at risk of developing tumor
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -

Query Match   100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%;   Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

RESULT 4
AAB20085
ID   AAB20085 standard; Protein; 195 AA.
XX
AC   AAB20085;
XX
DT   23-APR-2001 (first entry)
XX
DE   Human CARD-5 protein.
XX
KW   CARD-5; caspase recruitment domain; human; cancer; infection;
KW   autoimmune disease; neurological disease; haematological disease;
KW   immune disease; inflammation; antitumour; antiseptic;
KW   immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Domain   111..181
FT   /note= "CARD"
XX
PN   WO200100826-A2.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US17691.
XX
PR   28-JUN-1999; 99US-0340620.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Bertin J;
XX
PS   Claim 9; Fig 21; 208pp; English.
XX
CC   The present sequence is that of human caspase recruitment domain 5
CC   (CARD-5), an intracellular protein predicted to be involved in
CC   regulating caspase activation. The sequence is predicted from an
CC   isolated cDNA clone (see AAF30007). Methods of diagnosing and
CC   treating patients suffering from a disorder associated with an
CC   abnormal level or rate of apoptotic cell death, abnormal activity
CC   of the Fas/APO-1 receptor complex, abnormal activity of the tumour
CC   necrosis factor receptor complex or abnormal activity of a caspase
CC   involve administering a compound that modulates the expression or
CC   activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
CC   therapy methods. Such disorders include cancer, viral infection,
CC   autoimmune disorders, neurological diseases, haematological
CC   disorders, inflammatory disorders and immune disorders. CARD-3,
CC   -4, -5 and -6 proteins can be used to regulate cell proliferation,
CC   cell survival and cell growth. They can also be used to screen
CC   drugs or compounds that modulate their activity or expression and
CC   to treat disorders associated with insufficient or excessive
CC   production of CARD-3, -4, -5 or -6 protein, or production of an
CC   aberrant protein.
XX
SQ   Sequence 195 AA;

DR   WPI; 2001-061973/07.
DR   N-PSDB; AAF30007.
XX
PT   Isolated intracellular proteins predicted to be involved in regulating
PT   caspase activation are used for diagnosis and treatment of e.g. cancer,
PT   viral infections, autoimmune diseases, neurological diseases and
PT   haematological disorders -

Query Match   100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%;   Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
```

```
Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
         |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

```
RESULT 5
AAU99353
ID   AAU99353 standard; Protein; 195 AA.
XX
AC   AAU99353;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human caspase recruitment domain-5 (CARD-5) protein.
XX
KW   Human; caspase recruitment domain-5; CARD-5; antiinflammatory;
KW   immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
KW   apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
KW   cell proliferation; gene therapy; immune disorder;
KW   chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
KW   sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
KW   human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
KW   lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
KW   arthritis; cell depletion; neurological disorder; Alzheimer's disease;
KW   Parkinson's disease; spinal muscular atrophy; haematologic disease;
KW   myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
KW   stroke.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          111..195
FT                   /label= CARD_domain
XX
PN   WO200244354-A2.
XX
PD   06-JUN-2002.
XX
PF   29-NOV-2001; 2001WO-US44894.
XX
PR   01-DEC-2000; 2000US-0728721.
PR   24-APR-2001; 2001US-0841879.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI; 2002-575538/59.
DR   N-PSDB; ABK87966.
XX
PT   Novel isolated murine or human caspase recruitment domain (CARD)-5
PT   polypeptide, useful for treating immune disorders such as Hashimoto's
PT   thyroiditis, graft rejection, allergy, glomerular nephritis,
PT   tuberculosis.
XX
PS   Claim 22; Fig 3; 100pp; English.
XX
CC   The invention discloses the isolated polypeptides, and encoding nucleic
CC   acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
CC   (cysteinyl aspartate-specific proteinases) are central to the apoptotic
CC   program and responsible for the degradation of cellular proteins that
CC   lead to the morphological changes seen in cells undergoing apoptosis.
CC   Caspases interact with other caspases via their CARDs and different
CC   subtypes of CARDs may confer binding specificity. CARD-5 is an
CC   intracellular protein that is predicted to be involved in regulating
CC   caspase activation. CARD-5 activates the nuclear factor-kappa B
CC   (NF-kappaB) transcription factor pathway and binds the CARDs of
CC   caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
CC   activity and NF-kappaB activation, regulate cell growth and cell death
CC   for identifying compounds which bind to them and modulate their activity and
CC   for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
CC   nucleic acids, antibodies and modulators of CARD-5 expression or activity
CC   can be used to treat immune disorders such as chronic inflammatory
CC   diseases and disorders, Hashimoto's thyroiditis, graft rejection,
CC   sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
CC   nephritis, human immunodeficiency virus (HIV) and bacterial infections
CC   (including tuberculosis and lepromatous leprosy) and in screening and
CC   detection assays. Modulators of CARD-5 activity or expression are also
CC   useful for treating autoimmune disorders and arthritis, cell depletion,
CC   erythematosus and arthritis, cell conditions, such as systemic lupus
CC   such as Alzheimer's disease, Parkinson's disease and spinal muscular
CC   atrophy, haematologic diseases, such as myelodysplastic syndrome and
CC   aplastic anaemia, myocardial infarction and stroke. The sequence
CC   presented is the human caspase recruitment domain-5 (CARD-5) protein.
XX
SQ   Sequence  195 AA;
```

```
Query Match             100.0%;  Score 990; DB 23;  Length 195;
Best Local Similarity   100.0%;  Pred. No. 3.2e-101;
Matches 195;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
         |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

```
RESULT 6
AAO17854
ID   AAO17854 standard; Protein; 195 AA.
XX
AC   AAO17854;
XX
DT   20-AUG-2002  (first entry)
XX
DE   Pyrin domain containing protein Pycard.
XX
KW   Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW   antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW   neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW   nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW   inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW   arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW   osteoarthritis; glomerulonephritis.
XX
OS   Unidentified.
XX
PN   WO200240668-A2.
XX
PD   23-MAY-2002.
XX
PF   30-OCT-2001; 2001WO-EP12545.
XX
PR   15-NOV-2000; 2000DB-1056687.
PR   30-NOV-2000; 2000DB-1059595.
```

```
XX
PA   (APOT-) APOTECH RES & DEV LTD.
XX
PI   Tschopp J, Martinon F;
XX
DR   WPI; 2002-427093/45.
DR   N-PSDB; AAL47126.
XX
PN
XX
PT   New DNA encoding protein with pyrin domain, useful for treating
PT   diseases involving impaired signal transduction, particularly
PT   inflammation, also proteins and antibodies -
XX
PS   Claim 5; Fig 1; 116pp; German.
XX
CC   The present invention relates the DNA and their encoded proteins, where
CC   the proteins contain at least one PYD (pyrin) domain. These can be used
CC   to treat diseases associated with impaired intracellular signal
CC   transduction, particularly inflammation such as psoriasis,
CC   arteriosclerosis, bacterial or viral infections (particularly meningitis
CC   and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC   sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC   and Parkinson's diseases. The present sequence is a protein of the
CC   invention.
XX
SQ   Sequence 195 AA;

Query Match             100.0%;  Score 990;  DB 23;  Length 195;
Best Local Similarity   100.0%;  Pred. No. 3.2e-101;
Matches  195;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRARDAILDALENLITAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRARDAILDALENLITAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy         61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy        121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSPTPAWWTCKDLLLQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSPTPAWWTCKDLLLQA 180

Qy        181 LRESQSYLVEDLERS 195
              |||||||||||||||
Db        181 LRESQSYLVEDLERS 195
```

```
RESULT 7
ABG71635
ID   ABG71635 standard; Protein; 195 AA.
AC   ABG71635;
XX
DT   10-JAN-2003  (first entry)
XX
DE   Human caspase recruitment domain-5 (CARD-5).
XX
KW   Human; caspase activity; caspase recruitment domain-5; CARD-5;
KW   caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
KW   pseudoICE; ICEBERG; cell growth; cell death; inflammation; CARD-7;
KW   apoptosis; caspase activation; cancer; follicular lymphoma;
KW   leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
KW   autoimmune disease; systemic lupus erythematosus; reactive arthritis;
KW   human immunodeficiency virus infection; HIV infection; ALS;
KW   Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW   myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
KW   insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
KW   graft rejection; allergic rhinitis; food allergy; conjunctivitis;
KW   glomerular nephritis; cytostatic; virucide; immunosuppressive;
KW   dermatological; nephrotropic; neuroprotective; cardiant;
KW   caspase recruitment domain-7.
OS   Homo sapiens.
```

```
XX
PN   US2002128198-A1.
XX
PD   12-SEP-2002.
XX
PF   27-NOV-2001; 2001US-0996617.
XX
DR   WPI; 2003-028967/02.
DR   N-PSDB; ABS56032.
XX
PI   Bertin J;
XX
PA   (BERT/) BERTIN J.
XX
PT   Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT   CARD-7 and CARD-5 in presence of test compound, measuring their
PT   binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PT   altered -
XX
PR   28-JUN-1999;  99US-0340620.
PR   27-OCT-1999;  99US-0428252.
PR   15-AUG-2001;  2001US-0931071.
XX
PS   Disclosure; Fig 7; 43pp; English.
XX
CC   The present invention relates to methods of identifying compounds
CC   that regulate caspase recruitment activity using caspase recruitment domain-7
CC   (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC   method for identifying a compound that modulates the interaction
CC   between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC   converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
CC   CARD-8 molecules are useful as modulating agents in regulating a
CC   variety of cellular processes including cell growth, cell death, and
CC   inflammation. The methods of the invention are useful for identifying
CC   compounds that have the ability to increase/decrease apoptosis, or
CC   comprise the ability to induce caspase activation. The methods are
CC   useful for treating a disorder associated with inappropriate apoptosis
CC   or inappropriate inflammation. The methods are useful for treating
CC   disorders associated with an undesirably low rate of apoptosis such
CC   as cancer (preferably follicular lymphoma, chronic myelogenous
CC   leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
CC   infections, autoimmune diseases caused by low levels of apoptosis
CC   (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC   and arthritis). The methods are also useful for treating disorders with
CC   undesirably high rates of apoptosis such as human immunodeficiency
CC   virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC   amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC   muscular atrophy, various forms of cerebellar degeneration, anaemia
CC   associated with chronic disease, aplastic anaemia, chronic neutropaenia,
CC   myelodysplastic syndromes, myocardial infarction, stroke, and
CC   various inflammatory disorders (e.g. Crohn's disease, reactive
CC   arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC   psoriasis, graft rejection, allergic rhinitis, food allergies,
CC   conjunctivitis, glomerular nephritis, etc). The present sequence
CC   represents human CARD-5. The sequence encoding CARD-5 is used to
CC   identify cDNA encoding CARD-7.
XX
SQ   Sequence 195 AA;

Query Match             100.0%;  Score 990;  DB 24;  Length 195;
Best Local Similarity   100.0%;  Pred. No. 3.2e-101;
Matches  195;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRARDAILDALENLITAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRARDAILDALENLITAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy         61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy        121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSPTPAWWTCKDLLLQA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   121  ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA  180

Qy   181  LRESQSYLVEDLERS  195
          |||||||||||||||
Db   181  LRESQSYLVEDLERS  195
```

RESULT 8
```
ID   AAG74647 standard; Protein; 205 AA.
XX
AC   AAG74647;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen protein SEQ ID NO:5411.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;  99US-0157137.
PR   03-NOV-1999;  99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   N-PSDB; AAH34052.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
PS   Claim 11; Page 7035-7036; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the protein. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAH77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
SQ   Sequence  205 AA;

Qy   1   MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11  MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY  70
```

Query Match   100.0%; Score 990; DB 22; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-101;
Matches 195; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy   181  LRESQSYLVEDLERS  195
          |||||||||||||||
Db   191  LRESQSYLVEDLERS  205

Qy   121  ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131  ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA  190

Qy   61   LETYGAELTANVLRDMGLQEMAGQLQAAATHQGSGNAAPAGTQAPPQGSAAKPGLHFIDQHRA  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71   LETYGAELTANVLRDMGLQEMAGQLQAAATHQGSGNAAPAGTQAPPQGSAAKPGLHFIDQHRA  130
```

RESULT 9
```
ID   AAB00594 standard; Protein; 176 AA.
XX
AC   AAB00594;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Alternatively spliced form of human TMS1 protein (lacking exon2).
XX
KW   Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast.
XX
OS   Homo sapiens.
XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999;  99US-0159975.
XX
PA   (UIEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI; 2001-290922/30.
DR   N-PSDB; AAD03906.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation of a CpG island in its 5' regulatory region. TMS1
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -
PS   Claim 85; Page 123; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it.
CC   The present sequence is alternatively spliced form of human target of
CC   methylation-induced silencing-1 (TMS1) protein lacking exon2.
SQ   Sequence  176 AA;
```

Query Match   88.2%; Score 873.5; DB 22; Length 176;
Best Local Similarity 90.3%; Pred. No. 2.3e-88;
Matches 176; Conservative   0; Mismatches   0; Indels 19; Gaps   1;

RESULT 10

```
ID   AAB43675 standard; Protein; 190 AA.
XX
AC   AAB43675;
XX
DT   08-FEB-2001 (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID NO:1120.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening.
XX
OS   Homo sapiens.
XX
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999; 99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   N-PSDB; AAC77884.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 11; Page 1731-1732; 2352pp; English.
XX
CC   AAC77607 to AAC77848 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 190 AA;
```

```
Query Match    76.7%;  Score 759;  DB 21;  Length 190;
Best Local Similarity   95.6%;  Pred. No. 1.2e-75;
Matches  153;  Conservative   3;  Mismatches   2;  Indels   2;  Gaps   1;
```

```
Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGRAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |||||||||||||||||||||||||||||||                ||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQ------------GLHFIDQHRA 101

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  102 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLLQA 161

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  162 LRESQSYLVEDLERS 176
```

RESULT 11

```
ID   AAE00592 standard; Protein; 193 AA.
XX
AC   AAE00592;
XX
DT   02-JUL-2001 (first entry)
XX
DE   Mouse target of methylation-induced silencing-1 (TMS1) protein.
XX
KW   Mouse; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast.
XX
OS   Mus musculus.
XX
XX
PN   WO20012923S-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999; 99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI; 2001-290922/30.
DR   N-PSDB; AAD03904.
XX
PT   Novel gene TMS1, transcriptionally silenced due to increased
PT   methylation useful for identifying subject at risk of developing tumor
PT   characterized by abnormal methylation, for treating cancer by inducing
PT   apoptosis -
XX
PS   Claim 85; Page 120; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
```

```
Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGRAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGRAAPAGIQAPPQSAAKPGLHFIDQHRA 130

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR 160
        |||||||||||||||||||||||||||||| |.|:|
Db  131 ALIARVTNVEWLLDALYGKVLTDEQYQAVR--PSPPTQAR 168
```

CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is mouse target of methylation-induced silencing-1
CC (TMS1) protein.
XX
SQ Sequence 193 AA;

Query Match 69.1%; Score 684; DB 22; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-67;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy 1 MGRARDAILDALENLITABELKKPFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
      |||||||||||||||:|:||||||||||||||||:|||||||||||||||||||||||||:|
Db 1 MGRARDAILDALENLSGDELKKPKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

Qy 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
       |||||||:|:||||:||||||||||||||||||||||||||||||||||||||||||||
Db 61 LESYGLEITMTVLRDMGLQELABQLQ-TTKEBSGAVAAAASVPAQSTARTG-HFVDQHRQ 118

Qy 121 ALIARVTNVEWLLDALYGKVLLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180
        ||||:|||||:|||||||||:||||||||||||||||||||||||||||:|:|||||||:
Db 119 ALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLLQA 178

Qy 181 LRESQSYLVEDLERS 195
        |:| |||||||:|
Db 179 LKEIHPYLVMDLEQS 193

RESULT 12
ID AAB20086 standard; Protein; 193 AA.
XX
AC AAB20086;
XX
DT 23-APR-2001 (first entry)
XX
DE Mouse CARD-5 protein.
XX
CC CARD-5; caspase recruitment domain; mouse; cancer; infection;
CC autoimmune disease; neurological disease; haematological disease;
CC immune disease; inflammation; antitumour; antiseptic;
CC immunomodulator; antinflammatory; apoptosis; diagnosis;
CC gene therapy.
XX
OS Mus sp.
XX
FH Key   Location/Qualifiers
FT Domain   110..179
FT   /note= "CARD"
XX
PN WO20010082Lang6-A2.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17691.
XX
PR 28-JUN-1999; 99US-0340620.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-061973/07.
DR N-PSDB; AAF33008.
XX
PT Isolated intracellular proteins predicted to be involved in regulating
PT caspase activation are used for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PT haematological disorders -

XX
PS Claim 9; Fig 19; 208pp; English.
XX
CC The present sequence is that of mouse caspase recruitment domain 5
CC (CARD-5), an intracellular protein predicted to be involved in
CC regulating caspase activation. The sequence is predicted from an
CC isolated cDNA clone (see AAF30008). It shows 71.8% amino acid
CC identity to human CARD-5 (see AAB20085). Methods of diagnosing
CC and treating patients suffering from a disorder associated with an
CC abnormal level or rate of apoptotic cell death, abnormal activity
CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
CC necrosis factor receptor complex or abnormal cell death, which
CC involve administering a compound that modulates the expression or
CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
CC therapy methods. Such disorders include cancer, viral infection,
CC autoimmune disorders, inflammatory disorders and immune disorders,
CC disorders, inflammatory disorders and immune disorders. CARD-3,
CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
CC cell survival and cell growth. They can also be used to screen
CC drugs or compounds that modulate their activity or expression and
CC to treat disorders associated with insufficient or excessive
CC production of CARD-3, -4, -5 or -6 protein, or production of an
CC aberrant protein.
XX
SQ Sequence 193 AA;

Query Match 69.1%; Score 684; DB 22; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-67;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy 1 MGRARDAILDALENLITABELKKPFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
      |||||||||||||||:|:||||||||||||||||:|||||||||||||||||||||||||:|
Db 1 MGRARDAILDALENLSGDELKKPKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

Qy 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
       |||||||:|:||||:||||||||||||||||||||||||||||||||||||||||||||
Db 61 LESYGLEITMTVLRDMGLQELABQLQ-TTKEBSGAVAAAASVPAQSTARTG-HFVDQHRQ 118

Qy 121 ALIARVTNVEWLLDALYGKVLLTDEQYQAVRAEPTNPSKQMRKLFSFTPAWNWTCKDLLLQA 180
        ||||:|||||:|||||||||:||||||||||||||||||||||||||||:|:|||||||:
Db 119 ALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLLQA 178

Qy 181 LRESQSYLVEDLERS 195
        |:| |||||||:|
Db 179 LKEIHPYLVMDLEQS 193

RESULT 13
ID AAU99352 standard; Protein; 193 AA.
XX
AC AAU99352;
XX
DT 07-OCT-2002 (first entry)
XX
DE Mouse caspase recruitment domain-5 (CARD-5) protein.
XX
KW Mouse; caspase recruitment domain-5; CARD-5; antiinflammatory;
KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
KW cell proliferation; gene therapy; immune disorder;
KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
KW stroke.
XX
OS Mus sp.
XX
FH Key   Location/Qualifiers

```
FT   Domain          110..193
FT                   /label= CARD_domain
XX
PN   WO200244354-A2.
XX
PD   06-JUN-2002.
XX
PF   29-NOV-2001; 2001WO-US44894.
XX
PR   01-DEC-2000; 2000US-0728721.
PR   24-APR-2001; 2001US-0841879.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Bertin J;
XX
DR   WPI: 2002-557538/59.
DR   N-PSDB; ABK87964.
XX
PS   Claim 22; Fig 1; 100pp; English.
XX
PT   Novel isolated murine or human caspase recruitment domain (CARD)-5
PT   polypeptide, useful for treating immune disorders such as Hashimoto's
PT   thyroiditis, graft rejection, allergy, glomerular nephritis,
XX
CC   The invention discloses the isolated polypeptides, and encoding nucleic
CC   acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
CC   (cysteinyl aspartate-specific proteinases) are central to the apoptotic
CC   program and responsible for the degradation of cellular proteins that
CC   lead to the morphological changes seen in cells undergoing apoptosis.
CC   Caspases interact with other caspases via their CARDs and different
CC   subtypes of CARDs may confer binding specificity. CARD-5 is an
CC   intracellular protein that is predicted to be involved in regulating
CC   caspase activation. CARD-5 activates the nuclear factor-kappa B
CC   (NF-kappaB) transcription factor pathway and binds the CARDs of
CC   caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
CC   activity and NF-kappaB activation, regulate cell growth and cell death
CC   and be used in gene therapy. The CARD-5 polypeptides are useful for
CC   identifying compounds which bind to them and modulate their activity and
CC   for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
CC   nucleic acids, antibodies and modulators of CARD-5 expression or activity
CC   can be used to treat immune disorders such as chronic inflammatory
CC   diseases and disorders, Hashimoto's thyroiditis, graft rejection,
CC   sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
CC   nephritis, human immunodeficiency virus (HIV) and bacterial infections
CC   (including tuberculosis and lepromatous leprosy) and in screening and
CC   detection assays. Modulators of CARD-5 activity or expression are also
CC   useful for treating autoimmune disorders, such as systemic lupus
CC   erythematosus and arthritis, cell depletion, neurological disorders,
CC   such as Alzheimer's disease, Parkinson's disease and spinal muscular
CC   atrophy, haematologic diseases, such as myelodysplastic syndrome and
CC   aplastic anaemia, myocardial infarction and stroke. The sequence
CC   presented is the mouse caspase recruitment domain-5 (CARD-5) protein.
XX
SQ   Sequence      193 AA;

Query Match       69.1%;  Score 684;  DB 23;  Length 193;
Best Local Similarity   71.8%;  Pred. No. 2.6e-67;
Matches 140;  Conservative 19;  Mismatches 34;  Indels 2;  Gaps 2;

Qy    1  MGRARDAILDALENLITAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY   60
            ||||||||||||||  :||||||||   ||||||||||||||||| ||| |||||||| :
Db    1  MGRARDAILDALENLSGDELKKFRMKLITVQLREGYGRIPRGALLQMDAIDLTDKLVSSY   60

Qy   61  LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGTQAPPQSAAKPGLHFIDQHRA  120
            ||:||:||| |||||||||  | ||| :       | |   ||| ||::| ||::|||:
Db   61  LESYGLELTMTVLRDMGLQOELABQLQ-TYKBESGAVAAASVPAQSTARG-HFVDQHRQ   118

Qy  121  ALIARVTNVEWLLDLALYGKVLTDEGYQAVRAEPTNPSKORKLFSFTPANVWTCKDLLLQA  180
            ||||||| |  ||||||:| ||| ||||||||| |  | :||||||:|  |:||||:|||||
Db  119  ALIARVTEVDGVLDLALHGSVLTEGGYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLLQA  178
```

```
Qy  181  LRESQSYLVEDLERS  195
            |:| ||| |||:|
Db  179  LKEIHPYLVMDLEQS  193

RESULT 14
AAE00593
ID   AAE00593 standard; Protein; 171 AA.
XX
AC   AAE00593;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Rat target of methylation-induced silencing-1 (TMS1) partial protein.
XX
KW   Rat; target of methylation-induced silencing-1; TMS1; cytostatic;
KW   antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW   caspase-recruiting domain; CARD; cancer; breast.
XX
OS   Rattus norvegicus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 36
FT                   /note= "Encoded by AGY"
XX
PN   WO200129235-A2.
XX
PD   26-APR-2001.
XX
PF   18-OCT-2000; 2000WO-US28747.
XX
PR   18-OCT-1999;  99US-0159975.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Vertino PM;
XX
DR   WPI: 2001-290922/30.
DR   N-PSDB; AAD03905.
XX
PS   Claim 85; Page 121; 124pp; English.
XX
CC   The invention relates to identification of target of methylation-induced
CC   silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC   abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC   consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC   plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC   as tools for diagnosing and treating a subject at risk of developing
CC   cancer (e.g. breast cancer) characterised by abnormal methylation or
CC   abnormally low levels of TMS1 expression products. Unique fragments of
CC   TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC   TMS1 molecule is also useful for treating abnormal cell proliferation by
CC   increasing TMS1 polypeptide level to an above normal level. The CpG
CC   island region of TMS1 or its fragments are used to study the methylation
CC   patterns apart from any coding region contained in it.
CC   The present sequence is rat target of methylation-induced silencing-1
CC   (TMS1) partial protein.
XX
SQ   Sequence      171 AA;

Query Match       58.2%;  Score 576;  DB 22;  Length 171;
Best Local Similarity   67.1%;  Pred. No. 2e-55;
Matches 116;  Conservative 18;  Mismatches 37;  Indels 2;  Gaps 2;

Qy   23  FKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMA   82
            ||:||| :|:|||||||||||||| ||| ||||||||: :||:|||| ||:||||||:|| :
Db    1  FKIKLLTAPVREGYGRIPRGALLQMDPIDLTDKLVXYYLEGYGLELTMTVLRDMGIQELA   60
```

```
Qy     83  GQLQAATHGQSGNAAPAGIQAPPQSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLT 142
               |||:|   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  EQLQKIMEE-SGAVATATSVPAQGTARTE-HFVDQHRQALIARVTEVDGLLDALYGNVLT 118

Qy    143  DEQYQAVRAEPTNPSKMRKLPSFTPAWNWTCKDLLLQALRESQSYLVEDLERS 195
               :|  |||   ||  ::|:|   ||||||  | : ||| : |||| : |:|
Db    119  EGQYQAVRAETTNQNKMRKLPSFAPAWNLTCKNLFLEALRQTQPYLVTDLEQS 171
```

```
AAG74648
ID   AAG74648 standard; Protein; 158 AA.
XX
AC   AAG74648;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen protein SEQ ID NO:5412.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;  99US-0157137.
PR   03-NOV-1999;  99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,   Birse CE,   Rosen CA;
XX
DR   WPI; 2001-233357/24.
DR   N-PSDB; AAH34053.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 11; Page 7036-7037; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37195 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence   158 AA;

Query Match             57.0%;  Score 564.5;  DB 22;  Length 158;
Best Local Similarity   75.8%;  Pred. No. 3.4e-54;
Matches  119;  Conservative   3;  Mismatches   2;  Indels  33;  Gaps   2;

Qy      4  ARDAILDALENLTABEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLET 63
               |||||||||||||||||||||||||||
```

Search completed: January 29, 2004, 13:45:54
Job time : 9.48584 secs

```
Db     13  ARDAILDALENLTABELKKF-------------------------------------KLVSFYLET 41
               ||||||||||||||||||                                       |||||||||
Qy     64  YGAELTANVLRDMGLQEMAGQLQAAATHQGSGNAAPAGIQAPPQSAAKPGLHFIDQHRAALI 123
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42  YGAELTANVLRDMGLQEMAGQLQAAATHQGSGNAAPAGIQAPPQSAAKPGLHFIDQHRAALI 101

Qy    124  ARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR 160
               ||||||||||||||||||||||:  |:| :||||||
Db    102  ARVTNVEWLLDALYGKVLTDBQYQAVR--PSPPTQAR 136
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: January 29, 2004, 13:44:42 ; Search time 4.80296 Seconds
(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTABEL.........LLIQALRESQSYIVEDLRS 195
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%
Listing first 45 summaries

Database :   1: pir1:*
             2: pir2:*
             3: pir3:*
             4: pir4:*
             PIR_76:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	250	25.3	1192	2	T17255	hypothetical prote
2	87	8.8	323	2	T49723	hypothetical prote
3	86.5	8.7	372	2	S23326	gene ML2.2 protein
4	86.5	8.7	439	2	S75545	hypothetical prote
5	85	8.6	237	2	E55210	heti protein - Ana
6	85	8.6	1488	2	AH0615	cell division prot
7	84	8.5	292	2	C64699	proteinase IV (EC
8	83.5	8.4	542	2	A84229	sensory histidine
9	83	8.4	455	2	JC6051	trypsin-like prote
10	83	8.4	545	2	A87136	hypothetical prote
11	83	8.4	877	2	JN0772	glucan endo-1,3-be
12	82.5	8.3	446	2	S16308	translation elonga
13	82.5	8.3	522	2	D82195	phage replication
14	82	8.3	292	2	F71819	probable proteinas
15	82	8.3	788	2	A72330	phenylalanine-tRNA
16	81	8.2	792	2	A80704	pyruvate, water di
17	80.5	8.1	718	2	B81058	guanosine-3',5'-bi
18	80.5	8.1	725	2	B81819	guanosine-3',5'-bi
19	80.5	8.1	2117	2	T36180	CDA peptide synthe
20	80	8.1	596	2	F87297	feruloyl-CoA synth
21	80	8.1	1233	2	T40059	chromosome segrega
22	80	8.1	350	2	E87714	hypothetical prote
23	79	8.0	455	2	C91142	serine endoprotein
24	79	8.0	455	2	F85987	serine endoprotein
25	79	8.0	855	2	G82310	alanyl-tRNA synthe
26	79	8.0	1486	2	C64832	cell division prot
27	79	8.0	1486	2	G90754	kinesin-like cell
28	79	8.0	1486	2	E85618	hypothetical prote
29	79	8.0	1534	2	JH0228	cell division prot
30	78.5	7.9	322	2	F87549	type IV secretion
31	78.5	7.9	346	2	A48470	translation elonga
32	78.5	7.9	424	2	A2975	oxidoreductase ord
33	78.5	7.9	424	2	H98307	probable oxidoredu
34	78.5	7.9	699	2	T18984	hypothetical prote
35	78.5	7.9	956	2	B71468	probable insulinas
36	78.5	7.9	968	2	T51933	kinesin motor prot
37	78.5	7.9	4684	2	A59404	plectin [imported]
38	78	7.9	325	2	A97482	hypothetical prote
39	78	7.9	325	2	A12699	conserved hypothet
40	78	7.9	335	2	C75619	histidinol-phospha
41	78	7.9	351	2	AF3419	phage host specifi
42	78	7.9	366	2	G84849	probable actin [im
43	77.5	7.8	228	2	C87467	ABC transporter, A
44	77.5	7.8	542	2	JN0661	heat shock protein
45	77.5	7.8	542	2	S32106	groEL protein - La

ALIGNMENTS

RESULT 1
T17255
hypothetical protein DKFZp58601822.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17255
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17255
A;Molecule type: mRNA
A;Status: preliminary
A;Cross-references: EMBL:AL117470
A;Experimental source: adult uterus; clone DKFZp58601822
C;Genetics:
A;Note: DKFZp58601822.1

Query Match 25.3%; Best Local Similarity 44.5%; Score 250; DB 2; Length 1192;
Matches 65; Conservative 23; Mismatches 48; Indels 10; Gaps 4; Pred. No. 1.7e-13;

Qy   54 DKLVS-FYLETYGAELTANV--LRDMGLQEMA----GQLQAATHQGSGAAPAGIQAPPQS 106
            :||:: ||:|| ::|: | : |:|: |||| :|| |:|:| | : |:|: |:
Db 1036 DQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPAT---TIPPARIAVPSPL 1092

Qy   107 AAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPFPNPSKWRKLFSFT 166
            :| ||||:||:| :|:|||| |:|:|| ||:||:::||: |:|: ||:|||||||||:
Db 1093 DAPQLLHFVDQYREQLIARVTSVEVVLDKIHGQVLSQEQYERVLAENTRPSQWRKLFSLS 1152

Qy   167 PAWWWTCKDLLLLQALRESQSYIVEDL 192
            ::| : ||| |:||| | :|: :|::|
Db 1153 QSWDRKCKDGLYQALKETHPHLIMEL 1178

RESULT 2
T49723
hypothetical protein B23L21.360 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49723
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49723
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <SCH>
A;Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.360
A;Experimental source: BAC clone B23L21; strain OR74A
C;Genetics:
A;Gene: NCSP:B23L21.360

A;Map position: 6
A;Introns: 183/1; 229/2

```
Query Match       8.8%;  Score 87;  DB 2;  Length 323;
Best Local Similarity 20.7%;  Pred. No. 5.1;
Matches 44;  Conservative 41;  Mismatches 80;  Indels 48;  Gaps 9;

Qy    6 DAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYG 65
Db  119 DNHFDLLKVQCGEDFEWFDEQVKSLP------------SSPVTIERFAFTENMAKHYIRSRG 168

Qy   66 AELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALIAR 125
Db  169 -----------EMSVR-MPGVVVA--HAGSGTILEVLRLQVPLVVVPNPTLMDNFHQAELT- 213

Qy  126 VTNVEWLLDALYGKV--LTDEQYQAVRAEPTNPSKMRKLFSFTP------------AWNW 171
Db  214 AQELESTGDAVYGRLGKLTBAIARSLELVAQGPLIKLKDLPPYSPPPFVPDSERVTLFDW 273

Qy  172 ----TC-----KDLILQALRESQSYLVEDLERS 195
Db  274 MTLTCYPDELRKQQHLQDLRNVEKSFQRRLEQA 306
```

RESULT 3
S23326
gene ML2.2 protein precursor - Streptococcus pyogenes
C;Species: Streptococcus pyogenes
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C;Accession: S23326
R;Bessen, D.E.; Fischetti, V.A.
Infect. Immun. 60, 124-135, 1992
A;Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A strep
A;Reference number: S23326
A;Accession: S23326
A;Molecule type: DNA
A;Status: preliminary
A;Residues: 1-372 <BES>
A;Cross-references: EMBL:X61276; NID:g47369; PIDN:CAA43582.1; PID:g47371
C;Superfamily: M5 protein

```
Query Match       8.7%;  Score 86.5;  DB 2;  Length 372;
Best Local Similarity 26.8%;  Pred. No. 6.7;
Matches 53;  Conservative 30;  Mismatches 78;  Indels 37;  Gaps 9;

Qy   12 LENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDAL-----DLTDKLVSFYLETYGA 66
Db   78 LEKINAEEBKNKLEAINKELNENYYKLQDG-----IDALBEKEDLKTTLAKTTKENEIS 133

Qy   67 ELTANVL-RDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQ--SAAKPGL-HFIDQHRAAL 122
Db  134 EASRKGLSRDLEASRTAKEBLEAAHQKLEAENKLTEGNQVSEASRKGLSNDLEASRAA- 192

Qy  123 IARVTNVEWLLDALYGKVLITDEQ-----YQAVRAEPTNPSKMRKLFSFTPAWWTCKDLL 177
Db  193 ------KKELEAXYQKLETDHQALEAKHQKLEADYQVSETSRKGLS------RD-- 234

Qy  178 LQALRESQSYLVEDLERS 195
Db  235 LEASREANKKVTSELTQA 252
```

RESULT 4
S75545
hypothetical protein slr0806 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
A;Accession: S75545
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o.K. Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75545
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <KAN>
A;Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAA18106.1; PID:g165531
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: glycolate oxidase chain glcD

```
Query Match       8.7%;  Score 86.5;  DB 2;  Length 439;
Best Local Similarity 24.0%;  Pred. No. 8.2;
Matches 42;  Conservative 22;  Mismatches 60;  Indels 51;  Gaps 8;

Qy   31 PLREGYGR---IPRGALLSMDALDLTDKLVSFYLETYGARLTANVLRDMGLQEMAGQLQA 87
Db  194 PVIHGYGTNGIIIEITLPLTPALPWREAIVSF---------------TNLSSAIFAFAQN 237

Qy   88 ATHQ--GSGAAPAGIQAPP------QSAAKPGLHFIDQHRAALIARVTNVEW----LL 133
Db  238 LAHQDGIVSKEISIQADPIPQYFSSLKSYQPGAHWV-------MVIVSELDWLAFTQLA 290

Qy  134 DALYGKVLITDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDLLILQALRESQSYL 188
Db  291 KASKGEIIFEQ---------DPQSPGKKINLIEF----NWWHTTLLARAVDSLTYL 334
```

RESULT 5
E55210
hetI protein - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Variety: PCC 7120
C;Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 15-Oct-1999
A;Accession: E55210
R;Black, T.A.; Wolk, C.P.
J. Bacteriol. 176, 2282-2292, 1994
A;Title: Analysis of a Het - mutation in Anabaena sp. strain PCC 7120 implicates a secon
A;Reference number: A55210; MUID:94209228; PMID:8157596
A;Accession: E55210
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <BLA>
A;Cross-references: GB:L22883; NID:g441119; PIDN:AAA22003.1; PID:g441121
C;Genetics:
A;Gene: hetI
A;Start codon: TTG

```
Query Match       8.6%;  Score 85;  DB 2;  Length 237;
Best Local Similarity 23.7%;  Pred. No. 5.2;
Matches 49;  Conservative 32;  Mismatches 72;  Indels 54;  Gaps 12;

Qy   14 NLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLV---SFYLETYGABLTA 70
Db   12 NLT---LLSDEVHLWRIPLDQ------PESQLQDLAATLSSDELARANRFYFPEHRRFTA 63

Qy   71 --NVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKP-----------GLHFIDQHR 119
Db   64 GRGIIRSI---LGGYL------GVEPQGVKPDYBSRGKPILGDRFAESGLLFNLSHS 111

Qy  120 AALIARVTN------------VEWL-----IDALYGKVLITDEQYQAVRAEPTNPSKMRKLFSFT 166
Db  112 QNLALCAVNYTRQIGIDLEYLRPTSDLESLAKRFPLPREYELLRSLP-DEQKQKIFFRY- 169

Qy  167 PAWWTCKDLLILQALRESQSYLVEDLE 193
Db  170 ----WTCKBAYLKATGDGIAKL-EEIE 191
```

RESULT 6
AH0615
cell division protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (st
C;Species: Salmonella enterica subsp. enterica serovar Typhi
C;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AH0615
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AH0615
A;Accession: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1488 <PAR>
A;Reference number: AH0615
A;Cross-references: GB:AL513382; PIDN:CAD05394.1; PID:gi6502156; GSPDB:GN00176
C;Genetics:
C;Gene: mukB
C;Superfamily: Escherichia coli mukB protein

Query Match   8.6%;   Score 85; DB 2; Length 1488;
Best Local Similarity 24.0%;
Matches   42; Conservative   33; Mismatches   56; Indels   44; Gaps   11;

Qy   3 RARDAILD-------ALENLTABELKKFKLKLLSVP--------LREGYGRIPRGALLS-   46
            |:|::|:: || || | : |||||:| |::| |::||||||
Db   638 RERBAIVERDEVGARKNAVDEEIER-----LSQPGGAEDQRLNALAERFG----GVLLSE 688

Qy   47 -MDALDLITDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAG--IQA 102
            || || | : : :||| |:| ||||||| ||| |||||| | |:| | ||
Db   689 IYDDVSLED---APYFSALYGPSRHAIVVPD--LSQIAEQLEGLTD----CPEDLYLIEG 739

Qy   103 PPQSAAKPGLHFIDQHRRALIARVTNVEWLLD------ALYGKVLTDEQYQAVARAB   152
            ||||| :|| | |||: |||:|||||| ||| |||| |||||| ||||
Db   740 DPQS-FDDSVFSVDELEKAVVKIADRQWRYSRFPSLPIFGRAARENRIESLHAE   793

RESULT 7
C64699
proteinase IV (EC 3.4.-.-) - Helicobacter pylori
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 15-Aug-1997 #text_change 29-Sep-1999
C;Accession: C64699
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.;
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: C64699
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-292 <TOM>
A;Cross-references: GB:AE000644; GB:AE000511; NID:g2314609; NID:g2314609; PIDN:AAD08477.1; PID:g2314614
C;Superfamily: proteinase IV MTH806
C;Keywords: hydrolase

Query Match   8.5%;   Score 84; DB 2; Length 292;
Best Local Similarity 23.6%;
Matches   30; Conservative   41; Mismatches   61; Indels   42; Gaps   7;

Qy   42 GALLSMDALDL-TDKL-------------VSFYLETYGAELTANV-----LRDMGLQEMAGQLQA   87
            ||||||| : || |:| ||||||||:: | ||:||||||| |
Db   57 GAIFSTEDFDKEVDKILKTPSIKGVLLLIDSPGGAVSASVELSEKIADL-KQXMPVLAYA 115

Qy   88 ATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRRALIARV-----TNVEWLLDALYGKVL 141
            | | | |:| | ||| | ||| | |||: |||:|| ||| ||||||| |
Db   116 RGVWASGSYYAGMQASBVVASK-----------ASLIGSIGVIFSGANVENLLNKV---- 160

Qy   142 TDEQYQAVRAEPTNPSKRMRKLFSTPAWWTCKDLLLQALRESQSYLVEDLERS 195
            |||||| |: || |: | ||: |||: | ::||| |:|| | ||
Db   161 ------GVATQGVHAGEYKEIGTFFRAWKPNEKDFLQNLVNEQYQMFVNDVAKA 208

RESULT 8

A84229
sensory histidine protein kinase homolog [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001
C;Accession: A84229
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84229
A;Status: preliminary
A;Residues: 1-542 <STO>
A;Cross-references: GB:AE004437; NID:g10580295; PIDN:AAG19197.1; GSPDB:GN00138
C;Genetics:
C;Gene: afsQ2

Query Match   8.4%;   Score 83.5; DB 2; Length 542;
Best Local Similarity 28.2%;
Matches   37; Conservative   18; Mismatches   53; Indels   23; Gaps   6;

Qy   37 GRIPRGALLSMDAL--DLTDKLVSFYLE----TYGAELTANVLRDM--GLQEMAGQLQAA   88
            ||| ||||||||| ||||||| || : ||||||| || | ||||||||| |
Db   40 GRIEYANEAARDVIGADARDLVGTFPWQAPWFTHDQDVQAAVREQVAAAVQGEASQF-TA   98

Qy   89 THQGSGAAPAGIQAPPQSAAKPGLHFIDQH-------------RAALIARVTNVEWLLD--- 134
            | | | | | | |||| ||||| | ||| |||:|||||| |||
Db   99 THRISGGGGTATVELELQPMPAPAVDGIDDHDHAFAVVVGRRAASESTAADVESALDAVQ 158

Qy   135 ALYGKVLTDEQ 145
            ||| : |:|
Db   159 ALYAATDTDDQ 169

RESULT 9

JC6051
trypsin-like proteinase (EC 3.4.21.-) hhoA precursor - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 31-Dec-1976 #sequence_revision 31-Dec-1996 #text_change 01-Mar-2002
C;Accession: JC6051; D65115
R;Bass, S.; Gu, Q.; Christen, A.
J. Bacteriol. 178, 1154-1161, 1996
A;Title: Multicopy suppressors of Prc mutant Escherichia coli include two HtrA (DegP) pro
A;Reference number: JC6051; MUID:96165273; PMID:8576052
A;Accession: JC6051
A;Molecule type: DNA
A;Residues: 1-455 <BAS>
A;Cross-references: GB:U15661; NID:g558911; PIDN:AAC43992.1; PID:g558912
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D65115
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-455 <BLAT>
A;Cross-references: GB:AE000402; GB:U00096; NID:g1789619; PIDN:AAC76266.1; PID:g1789629;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
C;Gene: hhoA
C;Superfamily: Helicobacter serine proteinase
C;Keywords: hydrolase; serine proteinase
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-455/Product: trypsin-like serine protease hhoA #status predicted <MAT>
F;109,139,214/Active site: His, Asp, Ser #status predicted

Query Match   8.4%;   Score 83; DB 2; Length 455;
Best Local Similarity 21.9%;
Matches   55; Conservative   34; Mismatches   76; Indels   86; Gaps   10;

```
Qy   11 ALENLTABELKKFKLKLLSVPLREGYGRIPRG-ALLSMDALDLTDKLVSF------YLET 63
Db  217 ALLNLNGB-------LIGINTAILAPGGSVGIGFAIPSNMARTLAQOLIDFGEIKRGLLGI 271

Qy   64 YGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQA---------------------- 102
Db  272 KGTEMSADIAKAFNLDVORGAFVSEVLPGSGSAKAGVRAGDIITSLNGKPLNSPAELRSR 331

Qy  103 ----PPQSAAKPGLHFIDQHRAALIARVT--------NVEWLLDALYGKVLTDEQY--- 146
Db  332 IATTEPGTKVKLGLL---RNGKPLEVEVTLIDTSTSSSASAEMITPALLEGATLSDGQLKDG 388

Qy  147 -------QAVRAEP-----------------TNPSKMRKLFSFTPAWNWTCKDLLLQ 179
Db  389 GKGIKIDEVVKGSSPAAQAGLQKDDVIIGVNRDRVNSIABMRKVLAAKPA------IIALQ 442

Qy  180 ALRESQS-YLV 189
Db  443 IVRGNESIYLL 453
```

RESULT 10
A87136
hypothetical protein ML1816 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87136
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R;Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: A87136
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-545 <STO>
C;Genetics:
C;Cross-references: GB:AL450380; NID:g13093526; PIDN:CAC30769.1; GSPDB:GN00147
A;Gene: ML1816
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

```
Query Match             8.4%;   Score 83;   DB 2;   Length 545;
Best Local Similarity  24.4%;   Pred. No. 22;
Matches   41;   Conservative 15;   Mismatches  50;   Indels  62;   Gaps  7;

Qy   51 DLTDKLVSFY--------LETYGAELTA-NVLRDMGLQ-----------------------EM 81
Db  113 DARDRAIRRYGOLEERFLALGGYSAESEAGRICASLGLPERVLVQQLCTLSGGQRRVEL 172

Qy   82 AGOLQAATHQGSGAAPAG------IQAPPQSAAKPGLHF--------------IDQHRAALIA 124
Db  173 ARILPAASRAGTGASGSGCTTLLLIDEPTNHLDADSLGWIRDFLRSHTGGLVVISHNVELIA 232

Qy  125 RVTNVEWLLDALYGKV-----------LTDEQYQAVRAEPTNPSK 158
Db  233 AVVNRWFLDAVLGKVDVYNMGWYKYLDSRATDEQRR--RRERVNAER 278
```

RESULT 11
JN0772
glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) precursor - Bacillus circulans (strain
N;Alternate names: beta-1,3-glucanase
C;Species: Bacillus circulans
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C;Accession: JN0772; PN0613
R;Yamamoto, M.; Aono, R.; Horikoshi, K.
Biosci. Biotechnol. Biochem. 57, 1518-1525, 1993
A;Title: Structure of the 87-kDa beta-1,3-glucanase gene of Bacillus circulans IAM1165 a
A;Reference number: JN0772; MUID:94033858; PMID:7764221
A;Accession: JN0772
A;Molecule type: DNA
A;Residues: 1-877 <YAM>
A;Cross-references: GB:D17519; NID:g601877; PIDN:BAA04469.1; PID:g601878
A;Accession: PN0613
A;Molecule type: protein
A;Residues: 39-58 <YA>
A;Genetics:
A;Gene: bglH
A;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-38/Domain: signal sequence #status predicted <SIG>
F;39-877/Product: glucan endo-1,3-beta-D-glucosidase #status predicted <MAT>

```
Query Match             8.4%;   Score 83;   DB 2;   Length 877;
Best Local Similarity  22.3%;   Pred. No. 39;
Matches   40;   Conservative 26;   Mismatches  73;   Indels  40;   Gaps  8;

Qy   10 DALENLTABELKKFKLKLLSVPLR-----BGYGRIPRGALLSMDALDLTDKLVSFYLE-- 62
Db  471 DRAQNFVQD-GKLNIKALNEPKSFPQDPSRYAQYSSGKINTKDHFSLKYGRVDFRAKLP 529

Qy   63 ----TYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQS 106
Db  530 TGNGIWPALWMLPQDNVYGTWASSG---EIDWMEAKGRLPGST--SGAVHFGGQWPTNR 583

Qy  107 AAKPGLHF------IDQHRAALIARVTNVEWLLDA-LYGKVLTDEQYQAVRAEPTNPS 157
Db  584 YLSGEYHFPEGQTFANDVHVYSVVWEEDNIKWVVDGKFFFKVTRDQWYSA--AAPNPN 640
```

RESULT 12
S16308
translation elongation factor eEF-1 alpha chain - Stylonychia lemnae
C;Species: Stylonychia lemnae
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
C;Keywords: GTP binding; nucleotide binding; protein biosynthesis
C;Accession: S16308
R;Bierbaum, P.; Doenhoff, T.; Klein, A.
Mol. Microbiol. 5, 1567-1575, 1991
A;Title: Macronuclear and micronuclear configurations of a gene encoding the protein syn
A;Reference number: S16308; MUID:92157884; PMID:1840642
A;Accession: S16308
A;Molecule type: DNA
A;Residues: 1-446 <BIE>
A;Cross-references: EMBL:X57926; NID:g10139; PIDN:CAA41001.1; PID:g10140
A;Note: the authors translated the codon GAT for residue 307 as Gly
A;Genetics:
A;Gene: efaA
A;Genetic code: SGC5
C;Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
C;Keywords: GTP binding; nucleotide binding; protein biosynthesis
F;8-156/Domain: translation elongation factor Tu homology <ETU>
F;14-21/Region: nucleotide-binding motif A (P-loop)
F;153-156/Region: GTP-binding NKXD motif

```
Query Match             8.3%;   Score 82.5;   DB 2;   Length 446;
Best Local Similarity  21.4%;   Pred. No. 19;
Matches   44;   Conservative 40;   Mismatches  81;   Indels  41;   Gaps  8;

Qy    8 ILDALENLTAEELKKFKLKLLSVPLREGY------------GRIPRGAL----LSMDALDL 52
Db  216 LIDALDAL--DQPKRPKDKPLRLPLQDVYKIGGIGTVPVGRVETGLLKPGMVLTFAPMNI 273

Qy   53 TDKLVSFYLETYGABLTA------NV---LRDMGLQEMAGQLQAATHQGSGA------AP 97
Db  274 TTECKS--VEMHHESLITERPGDNVGFTVKNLSVKDLRRGYVASDSKNDPAKDTTNFLAQ 331

Qy   98 AGIQAPPQSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDE------QYAVR 150
Db  332 VIVLNHPGQIQKGYAPVLDCHTAHIACKFPEIESKVDRRSGKVLEEBKFIKSGEAALVR 391

Qy  151 AEPTNPSKMRLFSFTPAWNWTCKDL 176
Db  392 MVPQKPMCVEAFNQYPPLGRFAVRDM 417
```

RESULT 13
D82195
phage replication protein Cri VC1469 VC1475 [imported] - Vibrio cholerae (strain N16961)
C;Species: Vibrio cholerae
A;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: D82195; A82196
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-522 <HEI>
A;Cross-references: GB:AE004225; GB:AE003852; NID:g9655968; PIDN:AAP94626.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
A;Accession: A82196
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-522 <HE2>
A;Cross-references: GB:AE004225; NID:g9655968; PIDN:AAP94631.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics: <GEN1>
A;Gene: VC1469
A;Map position: 1
C;Genetics: <GEN2>
A;Gene: VC1475
A;Map position: 1

Query Match   24.4%;
Best Local Similarity   24.4%;   Score 82.5;   DB 2;   Length 522;
Matches   38;   Pred. No. 23;   Conservative   20;   Mismatches   53;   Indels   45;   Gaps   7;

Qy   10 DALENLTABELKKFKLKLLSVPLREGYGRI-PRGALLSMDALDLTDKLVSFYLETYGAEL   68
         || |||:: ||||:| | | |||||: ||:: |||||||||||||| ||||||| |:|
Db   300 DLLESLKGEELELYNDSKIKQLLRDAYSTMTPKGNISYAKA----DRLFRFYM------ 348

Qy   69 TANVLRDMGLQEMAGQLQAAT-HQGS-----GAAPAGIQAPPQSAAKPGLHFIDQHRAA 121
         |:|:| |: | | | :|| | ::| |||||| |||||||||:||
Db   349 ---TLCDRGYQELKAHSSKATLHRNMRDLMAIGFSKADLQNLSEGERMP-------- 394

Qy   122 LIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPS 157
         |: | : |:: |: |||||:|||||||||| 157
Db   395 -LAQVLNFNF---------DNQRPANYVEPVSPT 418

Matches   40;   Conservative   32;   Mismatches   60;   Indels   42;   Gaps   7;

Qy   42 GALLSMDALDL-TDKL---------VSFYLETYGAELTANV-----LRDMGLQEMAGQLQA   87
         ||: ||| || |:: || : :| :|| : |::: | :| :
Db   57 GAIFSTEDFDKEVDKILKTPSIKGVLLLIDSPGGAVSASVELSEKIADL-KQKMPVLAYA 115

Qy   88 ATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALIARV-------TNVEWLLDALYGKVL 141
         | ||| || ::| |:   :   ::| || ::   : |: :|| :
Db   116 RGVMASGSYYAGMQASEVYASK----------ASLLGSIGVIFSSANVENLLNKV----- 160

Qy   142 TDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQALRESQSYLVEDLERS 195
             |::| | | |::   : | |||: |:: || : | | |:| |
Db   161 ------GVARQGVHAGEYKEIGTFTRAWKPNEKEFLQNLVNEQYQMFVDDVAKA 208

RESULT 14
F71819
probable proteinase IV - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
A;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
C;Accession: F71819
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: F71819
A;Accession: F71819
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-292 <ARN>
A;Cross-references: GB:AE001556; GB:AE001439; NID:g4155938; PIDN:AAD06917.1; PID:g415559
A;Experimental source: strain J99
C;Genetics:
A;Gene: sppA
C;Superfamily: proteinase IV MTH006

Query Match   8.3%;   Score 82;   DB 2;   Length 292;
Best Local Similarity   23.0%;   Pred. No. 12;

RESULT 15
A72330
phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
A;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C;Accession: A72330
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: A72330
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-788 <ARN>
A;Cross-references: GB:AE001749; GB:AE000512; NID:g4981346; PIDN:AAD35904.1; PID:g4981135
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0822
C;Superfamily: phenylalanine-tRNA ligase beta chain
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match   8.3%;   Score 82;   DB 2;   Length 788;
Best Local Similarity   24.4%;   Score 82;   DB 2;   Length 788;
Matches   29;   Pred. No. 42;   Conservative   19;   Mismatches   57;   Indels   14;   Gaps   2;

Qy   6 DAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYG   65
         | :| |||| | |:|| || || | |: || :|: | | | |
Db   87 DYVILALEGATLN-----NGLKIEPRBFKGVISEGMLCSLEELGLBEEKSDRVRFPDP 139

Qy   66 AELTANVLRDMGLQEMAGQLQAATHQ------GSGAAPAGIQAPPQSAAKPGLHFIDQ 117
         | :| :   :| | : | :| |:| | |: |||| |: :|
Db   140 VELGVNVVEEYGLNERVLDIEITPNRPDCLSIIGVARELSALSGRPLNKPQPDVSFVDE 198

Search completed: January 29, 2004, 13:49:47
Job time : 6.80296 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 3.00185 seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-8

Sequence: 990

Perfect score: 990

Sequence: 1 MGRARDAILDALENLTAEEL.........LLLQALRESQSYLVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	990	100.0	195	1	ASC_HUMAN	Q9ULZ3 homo sapien
2	684	69.1	193	1	ASC_MOUSE	Q9EPB4 mus musculu
3	250	25.3	1473	1	NAL1_HUMAN	Q9C000 homo sapien
4	213	21.5	203	1	ASC_BRARE	Q91N6 brachydanio
5	138	13.9	854	1	PYA5_RAT	Q63035 rattus norv
6	132	13.3	843	1	PYA5_MOUSE	Q91W92 mus musculu
7	125.5	12.7	892	1	PYA5_HUMAN	P59044 homo sapien
8	122.5	12.4	781	1	MEFV_HUMAN	O15553 homo sapien
9	99.5	10.1	980	1	PYA3_HUMAN	Q8WX94 homo sapien
10	99	10.0	431	1	CAR8_HUMAN	Q9Y292 homo sapien
11	91.5	9.2	143	1	SILE_SALTY	Q94n3 salmonella
12	89	9.0	338	1	RUVB_THETN	Q8ran2 thermoanaer
13	86.5	8.7	372	1	M22_STRPY	P50469 streptococc
14	86.5	8.7	835	1	SYFB_CORGL	Q8ngm6 corynebacte
15	86	8.7	1033	1	CIS1_MOUSE	Q8r4b8 mus musculu
16	85	8.6	237	1	HET1_ANASP	P37695 anabaena sp
17	84.5	8.5	540	1	RIK2_HUMAN	O43353 h receptor-
18	83	8.4	378	1	ACT_SCHDU	O65314 scherffelia
19	83	8.4	455	1	DEGQ_ECOLI	P39099 escherichia
20	82.5	8.3	446	1	EF1A_STYLE	P25166 stylonychia
21	82	8.3	788	1	SYFB_THEMA	Q9wz89 thermotoga
22	80	8.1	539	1	RIK2_MOUSE	P58801 mus musculu
23	80	8.1	1034	1	CIS1_HUMAN	Q96p20 homo sapien
24	80	8.1	1233	1	SMC1_HUMAN	Q94383 schizosacch
25	79	8.0	860	1	SYA_VIBCH	Q56648 vibrio chol
26	79	8.0	1486	1	MUKB_ECOLI	P22523 escherichia
27	78.5	7.9	446	1	EF1A_EIMBO	Q07051 eimeria bov
28	78.5	7.9	4684	1	PLE1_HUMAN	Q15149 homo sapien
29	78	7.9	363	1	ACT9_ARATH	P93738 arabidopsis
30	76.5	7.7	213	1	Y914_THEMA	Q9X016 thermotoga
31	76.5	7.7	324	1	PARB_TREPA	O83295 treponema p
32	76.5	7.7	953	1	CAR4_HUMAN	Q9Y239 homo sapien
33	76	7.7	739	1	PURL_LISMO	Q8Y6G1 listeria mo
34	75.5	7.6	444	1	TIG_AZOBR	Q9x6w7 azospirillu
35	75.5	7.6	542	1	CH60_LACLA	P37282 lactococcus
36	75.5	7.6	542	1	CH60_LACLC	Q9aep7 lactococcus
37	75.5	7.6	546	1	CH60_LACAC	Q93g07 lactobacill
38	75.5	7.6	611	1	BIR_CHICK	Q90660 gallus gall
39	75.5	7.6	1698	1	Y076_HUMAN	O14999 homo sapien
40	75	7.6	545	1	CH60_PARDE	Q9z462 paracoccus
41	75	7.6	791	1	PPSA_ECOLI	P23538 escherichia
42	74.5	7.5	401	1	ISPE_LYCES	P93841 lycopersico
43	74.5	7.5	862	1	LOXA_PHAVU	P27480 phaseolus v
44	74.5	7.5	957	1	SECA_MYCSM	P71533 mycobacteri
45	74.5	7.5	1062	1	PYA7_HUMAN	P59046 homo sapien

ALIGNMENTS

RESULT 1
ASC_HUMAN

ID   ASC_HUMAN   STANDARD;   PRT;   195 AA.
AC   Q9ULZ3; Q96D12; Q9BS2S; Q9HBD0; Q9NXJ8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Apoptosis-associated speck-like protein containing a CARD (hASC)
DE   (PYCARD) (Target of methylation-induced silencing 1) (Caspase
DE   recruitment domain protein 5).
GN   ASC OR TMS1 OR CARD5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Leukemia;
RX   MEDLINE=20036508; PubMed=10567338;
RA   Masumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T.,
RA   Niikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.;
RT   "ASC, a novel 22-kDa protein, aggregates during apoptosis of human
RT   promyelocytic leukemia HL-60 cells.";
RL   J. Biol. Chem. 274:33835-33838(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Fibroblast;
RX   MEDLINE=20552139; PubMed=11103776;
RA   Conway K.E., McConnell B.B., Bowring C.E., Donald C.D., Warren S.T.,
RA   Vertino P.M.;
RT   "TMS1, a novel proapoptotic caspase recruitment domain protein, is a
RT   target of methylation-induced gene silencing in human breast
RT   cancers.";
RL   Cancer Res. 60:6236-6242(2000).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Martinon F., Hofmann K., Tschopp J.;
RT   "Pycard, a PYD and CARD containing molecule.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Bertin J.;
RT   "CARD5 protein is a CARD/PYRIN family member that is involved in
RT   apoptosis signal transduction.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Colon mucosa;
RA   Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA   Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA   Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RC   TISSUE=Lymph, and Pancreas;

RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RX   MEDLINE=20552140; PubMed=11103777;
RA   McConnell B.B., Vertino P.M.;
RT   "Activation of a caspase-9-mediated apoptotic pathway by subcellular
RT   redistribution of the novel caspase recruitment domain protein TMS1.";
RL   Cancer Res. 60:6243-6247(2000).
CC   -!- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
CC   activity is mediated predominantly through the activation of
CC   caspase 9.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
CC   a redistribution from the cytoplasm to the aggregates occurs.
CC   These appeared as hollow, perinuclear spherical, ball-like
CC   structures.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=3;
CC   Name=1;
CC   IsoId=Q9ULZ3-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=Q9ULZ3-2; Sequence=VSP_004119;
CC   Name=3;
CC   IsoId=Q9ULZ3-3; Sequence=VSP_004118;
CC   Note=Inferred from the cDNA sequence of Ref.5;
CC   -!- TISSUE SPECIFICITY: Widely expressed at low levels. Detected in
CC   peripheral blood leukocytes, lung, small intestine, spleen,
CC   thymus, colon and at lower levels in placenta, liver and kidney.
CC   Very low expression in skeletal muscle, heart and brain. Detected
CC   in the leukemia cell lines HL-60 and U937, but not in Jurkat T-
CC   cell lymphoma and Daudi Burkitt's lymphoma. Detected in the
CC   melanoma cell line WM35, but not in WM793 lymphoma. Detected in
CC   cervical carcinoma cells and Molt 4 lymphocytic leukemia cells.
CC   -!- MISCELLANEOUS: In breast tumorigenesis, methylation-mediated
CC   silencing may affect genes and proteins that act as positive
CC   mediators of cell death.
CC   -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   -!- CAUTION: Ref.5 sequence differs from that shown due to a
CC   frameshift in position 4.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AB023416; BAA87339.2; -.
DR   EMBL; AF184072; AAG01187.1; -.
DR   EMBL; AF184073; AAG01188.1; -.
DR   EMBL; AF255794; AAF99665.1; -.

DR   EMBL; AF310103; AAG30286.1; -.
DR   EMBL; AF384665; AAK63850.1; -.
DR   EMBL; AK000211; BAA91012.1; ALT_FRAME.
DR   EMBL; BC004470; AAH04470.1; -.
DR   EMBL; BC013569; AAH13569.1; ALT_INIT.
MIM; 606838; -.
DR   GO; GO:0006917; P:induction of apoptosis; TAS.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   Apoptosis; Anti-oncogene; Alternative splicing.
DR   PROSITE; PS50209; CARD; 1.
DR   PROSITE; PS50824; DAPIN; 1.
KW   Apoptosis; Anti-oncogene; Alternative splicing.
FT   DOMAIN   1   91   DAPIN.
FT   DOMAIN   107   195   CARD.
FT   VARSPLIC   26   85   Missing (in isoform 3).
FT   /FTId=VSP_004118.
FT   VARSPLIC   93   111   Missing (in isoform 2).
FT   /FTId=VSP_004119.
SQ   SEQUENCE   195 AA; 21627 MW; 45598728658686F46A CRC64;
     MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY   60
     LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA   120
     ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQVRKLFSFTPAWWTCKDLLLQA   180
     LRESQSYLVEDLERS   195

Qy   Query Match   100.0%; Score 990; DB 1; Length 195;
     Best Local Similarity   100.0%;
     Matches 195; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQVRKLFSFTPAWWTCKDLLLQA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKQVRKLFSFTPAWWTCKDLLLQA 180

Qy   181 LRESQSYLVEDLERS 195
              ||||||||||||||
Db   181 LRESQSYLVEDLERS 195

RESULT 2
ID   ASC_MOUSE   STANDARD;   PRT;   193 AA.
AC   Q9EPB4; Q9D2W9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Apoptosis-associated speck-like protein containing a CARD (mASC)
DE   (PYCARD).
GN   ASC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Thymus;
RX   MEDLINE=20580347; PubMed=11139337;
RA   Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;
RT   "Murine ortholog of ASC, a CARD-containing protein, self-associates
RT   and exhibits restricted distribution in developing mouse embryos.";
RL   Exp. Cell Res. 262:128-133(2001).
RN   [2]
SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast tumor;
RA   Martinon F., Hofmann K., Tschopp J.;
RT   "Pycard a PYD and CARD containing molecule.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Pancreas, and Tongue;

RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
CC   activity is mediated predominantly through the activation of
CC   caspase 9 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
CC   a redistribution from the cytoplasm to the aggregates occurs.
CC   These appeared as hollow, perinuclear spherical, ball-like
CC   structures (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in small intestine, colon, thymus,
CC   spleen, brain, heart, skeletal muscle, kidney, lung and liver.
CC   -!- DEVELOPMENTAL STAGE: Strongly expressed at E9.5 day in the
CC   telencephalon, thalamic areas of the diencephalon, heart and
CC   liver.
CC   -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AB032249; BAB16609.1; -.
DR   EMBL; AF310104; AAG30287.1; -.
DR   EMBL; AK009852; BAB25643.1; -.
DR   EMBL; AK007742; BAB25229.1; -.
DR   EMBL; AK018682; BAB31341.1; -.

DR   EMBL; BC008252; AAH08252.1; -.
DR   MGD; MGI:1931465; Asc.
DR   GO; GO:0005829; C:cytosol; IDA.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50209; CARD; 1.
DR   PROSITE; PS50824; DAPIN; 1.
KW   Apoptosis; Anti-oncogene.
FT   DOMAIN   1   91   DAPIN.
FT   DOMAIN   105   193   CARD.
FT   CONFLICT   159   159   K -> E (IN REF. 3).
SQ   SEQUENCE   193 AA; 21458 MW; 2A48AA0194870831 CRC64;

Query Match   69.1%; Score 684; DB 1; Length 193;
Best Local Similarity   71.8%; Pred. No. 1.8e-52;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy   1 MGRARDAILDALENLTAEELKKFKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              ||||||||||||||||:.|||||||||||||||||:|||||||||||||||||||||:|
Db   1 MGRARDAILDALENLSGDBLKKFKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSY 60

Qy   61 LETYGAELITANVLRDMGLQEMWAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              |||:||||||.|.|||||||:||||||||||||||
Db   61 LESYGLETMTVLRDMGLQELAEOLQ-TTKEESGAVAAAASVPAOSTARTG-HFVDQHRO 118

Qy   121 ALIARVTNVEWLLDALIYGKVLTDEQYQAVRAEPTNPSKRMRKLFSFTPAWNWTCKDLLLQA 180
              ||||||:.|:.||||||||:|||:||||||:|||||
Db   119 ALIARVTEVDVGVLIDALHGSVLITEGQYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLLQA 178

Qy   181 LRESQSYLVEDLERS 195
              |:|.|||||||:|
Db   179 LKEIHPYLVMDLEQS 193

RESULT 3
NA11_HUMAN
ID   NA11_HUMAN   STANDARD;   PRT;   1473 AA.
AC   Q9C000; Q9BZZ8; Q9BZZ9; Q9HAV0; Q9UFT4; Q9Y2S0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-
DE   forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
DE   caspase recruitment domain 7) (Caspase recruitment domain
DE   NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
GN   NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=21169419; PubMed=11270363;
RA   Bertin J., DiStefano P.S.;
RT   "The PYRIN domain: a novel motif found in apoptosis and inflammation
RT   proteins.";
RL   Cell Death Differ. 7:1273-1274(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=21148093; PubMed=11250163;
RA   Martinon F., Hofmann K., Tschopp J.;
RT   "The pyrin domain: a possible member of the death domain-fold family
RT   implicated in apoptosis and inflammation.";
RL   Curr. Biol. 11:R118-R120(2001).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=21153743; PubMed=11076957;
RA   Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
RA   Shi M.M., Vincenz C., Ward P.A.;
RT   "Molecular cloning and characterization of DEFCAP-L and -S, two
RT   isoforms of a novel member of the mammalian Ced-4 family of apoptosis
RT   proteins.";

```
RL   J. Biol. Chem. 276:9230-9238(2001).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1, 2, 3 AND 4), AND PROTEIN INTERACTION.
RC   TISSUE=T-cell;
RX   MEDLINE=21153744; PubMed=11113115;
RA   Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
RA   Godzik A., Reed J.C.;
RT   "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
RT   c-dependent caspase activation and apoptosis.";
RL   J. Biol. Chem. 276:9239-9245(2001).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [6]
RP   SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
RC   TISSUE=Uterus;
RA   Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Able to form cytoplasmic structures termed death
CC       effector filaments. Enhances APAF1 and cytochrome c-dependent
CC       activation of pro-caspase-9 and consecutive apoptosis. Seems to
CC       bind ATP.
CC   -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
CC       and with APAF1 in a cytochrome c-inducible way leading to the
CC       formation of an apoptosome. This interaction may be ATP-dependent.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1; Synonyms=NAC beta, DEFCAP-L;
CC         IsoId=Q9C000-1; Sequence=Displayed;
CC       Name=2; Synonyms=NAC alpha, DEFCAP-S;
CC         IsoId=Q9C000-2; Sequence=VSP_004327;
CC       Name=3; Synonyms=NAC gamma;
CC         IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
CC       Name=4; Synonyms=NAC delta;
CC         IsoId=Q9C000-4; Sequence=VSP_004326;
CC   -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
CC       expressed in peripheral blood leukocytes, chronic myelogenous
CC       leukemia cell line K-562, followed by thymus, spleen and heart.
CC       Also detected in lung, placenta, small intestine, colon, kidney,
CC       liver and muscle.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
```

```
DR   EMBL; AF298548; AAG15254.1; -.
DR   EMBL; AF310105; AAG30288.1; -.
DR   EMBL; AF229059; AAK00748.1; -.
DR   EMBL; AF229060; AAK00749.1; -.
DR   EMBL; AF229061; AAK00750.1; -.
DR   EMBL; AF229062; AAK00751.1; -.
DR   EMBL; AB023143; BAA76770.1; -.
DR   EMBL; AL117470; CAB55945.1; -.
DR   PIR; T17255; T17255.
DR   HSSP; P13489; 1A4Y.
DR   MIM; 606636; -.
DR   GO; GO:0005622; C:intracellular; IC.
DR   GO; GO:0016505; F:apoptosis activator activity; NAS.
DR   GO; GO:0008656; F:caspase activator activity; NAS.
DR   GO; GO:0019999; F:enzyme binding activity; IPI.
DR   GO; GO:0006919; P:induction of apoptosis; NAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR000767; Disease_resist.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00560; LRR; 2.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PRINTS; PR00364; DISEASERSIST.
DR   PROSITE; PS02209; CARD; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW   Alternative splicing.
FT   DOMAIN         1      92    DAPIN.
FT   DOMAIN       328     637    NACHT.
FT   REPEAT       704     725    LRR 1.
FT   REPEAT       807     830    LRR 2.
FT   REPEAT       864     887    LRR 3.
FT   REPEAT       921     944    LRR 4.
FT   REPEAT       950     973    LRR 5.
FT   REPEAT      1199    1215    LRR 6.
FT   REPEAT      1216    1236    LRR 7.
FT   DOMAIN      1374    1463    CARD.
FT   NP_BIND      334     341    ATP (POTENTIAL).
FT   VARSPLIC     958     987    Missing (in isoform 3 and isoform 4).
FT                               /FTId=VSP_004326.
FT   VARSPLIC    1262    1305    Missing (in isoform 2 and isoform 3).
FT                               /FTId=VSP_004327.
FT   MUTAGEN      340     340    K->L: NO EFFECT.
FT   MUTAGEN      340     340    K->S: NO EFFECT.
FT   CONFLICT     155     155    L -> H (IN REF. 1).
FT   CONFLICT     246     246    T -> S (IN REF. 1).
FT   CONFLICT     782     782    T -> S (IN REF. 1).
FT   CONFLICT     878     878    T -> M (IN REF. 1).
FT   CONFLICT     995     995    T -> I (IN REF. 1).
FT   CONFLICT    1119    1119    M -> V (IN REF. 1).
FT   CONFLICT    1184    1184    M -> V (IN REF. 1 AND 6).
FT   CONFLICT    1241    1241    V -> L (IN REF. 1).
FT   CONFLICT    1366    1366    R -> C (IN REF. 1).
SQ   SEQUENCE   1473 AA;  165865 MW;  438FDCE45C2562D CRC64;
```

```
RESULT 4

ASC_BRARE
ID   ASC_BRARE      STANDARD;      PRT;    203 AA.
AC   Q9I5N6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Apoptosis-associated speck-like protein containing a CARD (PYCARD).
GN   ASC OR ASC1.

Query Match             25.3%;  Score 250;  DB 1;  Length 1473;
Best Local Similarity   44.5%;
Matches   65;  Conservative   23;  Mismatches   48;  Indels   10;  Gaps   4;
                                   Pred. No. 9.6e-14;

Qy    54 DKLVS-FYLETYGAELTANV--LRDMGLQEMA-----GQLQAATHQGSGAAPAGIQAPPQS 106
            :| | :::|  :|:  :|:              ||:|||||||||||||::||||||||
Db  1317 DQLFSEFYVGHLSSGIRLQVKDKDETLVWEALVKPGDLMPAT---TLIPPARIAVPSPL 1373

Qy   107 AAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFT 166
         :|| |||||:| ||||:||||:|:|:|::||:|:||::||::|  | ||||:||||||:
Db  1374 DAPQLIHFVDQYRRQLIARVTSVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLS 1433

Qy   167 PAWWNTCKDLLLQALRESQSYLVEDL 192
          :|:  |||| ::||::|::|:::|
Db  1434 QSWDRKCKDGLYQALKETHPHLIMEL 1459
```

First Entry

```
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20373792; PubMed=10917738;
RA   Inohara N., Nunez G.;
RT   "Genes with homology to mammalian apoptosis regulators identified in
RT   zebrafish.";
RL   Cell Death Differ. 7:509-510(2000).
CC   -!- FUNCTION: Promotes caspase-mediated apoptosis (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
KW   Apoptosis; Anti-oncogene.
DR   EMBL; AF231013; AAF66956.1; -.
DR   ZFIN; ZDB-GENE-000511-2; ascl.
DR   InterPro; IPR001315; CARD.
DR   InterPro; IPR004020; PAAD_DAPIN.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50209; CARD; 1.
DR   PROSITE; PS50824; DAPIN; 1.
FT   DOMAIN        1     91       DAPIN.
FT   DOMAIN      112    203       CARD.
SQ   SEQUENCE   203 AA;  22867 MW;  EF457179EB7A78A7 CRC64;
```

RESULT 5

```
ID   PYA5_RAT     STANDARD;      PRT;   854 AA.
AC   Q63035;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYRIN-containing APAF1-like protein 5-like (Angiotensin II/vasopressin
DE   receptor).
GN   PYPAF5 OR AVR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
```

```
Query Match             21.5%;  Score 213;  DB 1;  Length 203;
Best Local Similarity   32.7%;  Pred. No. 1.6e-11;
  Matches   65;  Mismatches   83;  Indels   14;  Gaps    6;

Qy     5 RDAILDALENLTAEELKPKLKLLSVPLREGYGRIPRGALLSM-DALDLTDKLVSFYLET 63
            ::|:| :||||:|:||:|     |:||:|        ||::|| |:| :|
Db     6 KEHLQEAFEDLGADNLRKFKSKLGD--RRQEPRVTKSAIEKLKDBIDLADLMVGVFTSK 62

Qy    64 YGAELTANVLRDMGLQEMAGQLQAATHQSSGAAPAGIQAPPQSA------AKPGLHFIDQ 117
Db    63 DAVSVTVEILRAIKCNAVADDLLRNTGQSESKGALSDESKCASSKAVSKVAFSKVNFIDE 122

Qy   118 HRAALIARVTNVEWLLDAL-YGKVLTDEQYQAVRAEPTNPSKMRKLFS--FTPAWNWTCK 174
Db   123 HWKELIDRVNNVDPILDILRQKKVVITNEDYCTIRNKETPQKKWRELLTGPITCAGN-KGK 181

Qy   175 DLILQALRESQSYLVEDLE 193
         ::|||||||:| |:||
Db   182 EVLYDALRESNKFLMDDLE 200
```

Second Entry

```
RN   [1]
RP   CONCEPTUAL TRANSLATION OF 1-343.
RA   Hinz U.;
RL   Unpublished observations (FEB-2003).
RN   [2]
RP   SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
RC   TISSUE=Kidney;
RX   MEDLINE=96071640; PubMed=7489366;
RA   Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
RT   "Identification of a novel dual angiotensin II/vasopressin receptor on
RT   the basis of molecular recognition theory.";
RL   Nat. Med. 1:1074-1081(1995).
RN   [3]
RP   VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
RX   MEDLINE=21980185; PubMed=11984003;
RA   Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT   "The dual Angii/AVP receptor gene N119S/C163R variant exhibits
RT   sodium-induced dysfunction and cosegregates with salt-sensitive
RT   hypertension in the Dahl salt-sensitive hypertensive rat model.";
RL   Mol. Med. 8:24-32(2002).
RN   [4]
RP   IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA   Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT   "Identification of mammalian orthologs associates PYPAF5 with distinct
RT   functional roles.";
RL   FEBS Lett. 538:173-177(2003).
CC   -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC       activation of NF-kappa-B (By similarity). Angiotensin II and
CC       vasopressin binding protein. May stimulate cAMP accumulation.
CC   -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
CC       levels in all tissues tested.
CC   -!- DISEASE: Defects in PYPAF5 may be a cause of salt-sensitive
CC       hypertension.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC   -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC       sequences, in analogy to ortholog sequences.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; M85183; AAA03623.1; ALT_INIT.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   Pfam; PF00560; LRR; 2.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; leucine-rich repeat; Repeat.
FT   DOMAIN        1    103       DAPIN.
FT   DOMAIN      168    484       NACHT.
FT   DOMAIN      564    604       ASP/GLU-RICH.
FT   DOMAIN      655    662       POLY-LYS.
FT   REPEAT      433    458       LRR 1.
FT   REPEAT      610    633       LRR 2.
FT   REPEAT      722    745       LRR 3.
FT   REPEAT      812    836       LRR 4.
FT   NP_BIND     174    181       ATP (POTENTIAL).
FT   VARIANT     492    492       N -> S (IN SALT-SENSITIVE HYPERTENSION).
FT   VARIANT     536    536       C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT   MUTAGEN     766    766       E->K: ABOLISHES ANGIOTENSIN II BINDING.
SQ   SEQUENCE   854 AA;  95292 MW;  D7BBD922D77B3734 CRC64;
```

```
Query Match             13.9%;  Score 138;  DB 1;  Length 854;
Best Local Similarity   41.0%;  Pred. No. 0.00027;
```

Matches 32; Conservative 17; Mismatches 25; Indels 4; Gaps 2;

```
Qy     4 ARDAILDALENLTAEELKKFKLLSVPLREGYGR-IPRGALLSMDALDLTDKLVSFYLE 62
         |::|:||:||||:.::|:|:|||:|||  |||.||:|:||||||||||||||||||
Db    15 ARELLLAALQDLSQEQLKRFRHKLRDAPLD---GRSIPWGRLEHSDAVDLTDKLIEFYAP 71
```

Db   72 EPAVDVTRKILKADIRD 89

RESULT 6
PYA5_MOUSE
ID   PYA5_MOUSE   STANDARD;   PRT;   843 AA.
AC   Q91WS2; Q8KOL4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYPAF5.
OS   Mus musculus (Mouse).
GN   PYPAF5.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   CONCEPTUAL TRANSLATION OF 1-287.
RA   Hinz U.;
RL   Unpublished observations (FEB-2003).
RN   [2]
RP   SEQUENCE OF 288-843 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA   Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT   "Identification of mammalian ortholog associates PYPAF5 with distinct
RT   functional roles.";
RL   FEBS Lett. 538:173-177(2003).
CC   -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC   activation of NF-kappa-B (By similarity).
CC   -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC   -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC   sequences, in analogy to ortholog sequences.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial

DR   EMBL; BC013519; AAH13519.1; -.
DR   EMBL; BC031139; AAH31139.1; ALT_INIT.
DR   MGD; MGI:2141990; Pypaf5.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   Pfam; PF00560; LRR; 1.
DR   PROSITE; PS50837; NACHT; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN   11   102   DAPIN.
FT   DOMAIN   168   485   NACHT.
FT   DOMAIN   569   585   ASP/GLU-RICH.
FT   DOMAIN   654   661   POLY-LYS.
FT   REPEAT   434   459   LRR 1.
FT   REPEAT   609   632   LRR 2.
FT   REPEAT   811   834   LRR 3.
FT   NP_BIND   174   181   ATP (POTENTIAL).
SQ   SEQUENCE   843 AA; 94592 MW; 35FB7A7766A47DB51 CRC64;
```

Query Match          13.3%;  Score 132;  DB 1;  Length 843;
Best Local Similarity  41.0%;  Pred. No. 0.00089;
Matches   32; Conservative  14; Mismatches  28; Indels  4; Gaps  2;

```
Qy     4 ARDAILDALENLTAEELKKFKLLSVPLREGYGR-IPRGALLSMDALDLTDKLVSFYLE 62
         |::|:||:||||:.::|:|:|||:|||  |||.||:|:||||||||||||||||||||
Db    15 ARELLMATLEELSQEQLKRFRHKLRDAPLD---GRSIPWGRLERSDAVDLVDKLIEFYEP 71
```

Db   72 VPAVEMTRQVLKRSDIRD 89

RESULT 7
PYA5_HUMAN
ID   PYA5_HUMAN   STANDARD;   PRT;   892 AA.
AC   P59044;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   PYPAF5.
OS   Homo sapiens (Human).
GN   PYPAF5.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=22275822; PubMed=12387869;
RA   Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA   Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA   DiStefano P.S., Bertin J.;
RT   "Functional screening of five PYPAF family members identifies PYPAF5
RT   as a novel regulator of NF-kappaB and caspase-1.";
RL   FEBS Lett. 530:73-78(2002).
CC   -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC   activation of NF-kappa-B.
CC   -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
CC   much lower levels in T-cells.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
```

```
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; AF479748; AAL87105.1; -.
DR   InterPro; IPR007091; LRR_RI.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00368; LRR_RI; 3.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN        1    103       DAPIN.
FT   DOMAIN      196    513       NACHT.
FT   DOMAIN      604    614       POLY-GLU.
FT   REPEAT      462    487       LRR 1.
FT   REPEAT      727    747       LRR 2.
FT   REPEAT      755    778       LRR 3.
FT   REPEAT      811    834       LRR 4.
FT   REPEAT      845    868       LRR 5.
FT   NP_BIND     202    209       ATP (POTENTIAL).
SQ   SEQUENCE   892 AA;  98733 MW;  4AA8D1FC766DDE9D CRC64;

Query Match             12.7%;  Score 125.5;  DB 1;  Length 892;
Best Local Similarity   28.4%;  Pred. No. 0.0035;
Matches   48;  Conservative   25;  Mismatches   75;  Indels   21;  Gaps   5;

Qy    4  ARDAILDALENLTAFBLKKFKLLSVPLREGYGR-IPRGALLSMDALDLTDKLVSFYLE    62
         ||:||||||:||:||||::|:|||:||||||||| :||:|||||:||||||:||:
Db   18  ARELLLAALEEELSQEOLKRFRHKLRDV---GPDGRSIPWGRLERADAVDLAEQLAQFYGP   74

Qy   63  TYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSRAAKPGLHFIDQH----  118
         ::||                                              ||| |
Db   75  EPALEVARKTLKRADARDVAAQLQERRLQRLGLGSGTLLSVSEYKKYREHVLQLHARVK   134

Qy  119  -RAALIARVTNVEWLLDALYGKVLIDEQYQAVR-----AEPTNPSKMRK  161
          ::|| |                                |:|   | :|
Db  135  ERNARSVKITK------RFTKLLIAPESAAPEEALGPAEBEPEPGRARR  176

RESULT 8
ID   MEFV_HUMAN   STANDARD;   PRT;   781 AA.
AC   O15553;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Pyrin (Marenostrin).
GN   MEFV OR MEF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS FMF ILE-680, VAL-694 AND ALA-726.
RC   TISSUE=Leukocyte;
RX   MEDLINE=97433089; PubMed=9288758;
RA   Aksentijevich I., Centola M., Deng Z., Sood R., Balow J.E. Jr.,
RA   Wood G., Zaks N., Mansfield E., Chen X., Eisenberg S., Vedula A.,
RA   Shafran N., Raben N., Pras E., Pras M., Kastner D.L., Blake T.,
RA   Baxevanis A.D., Robbins C., Krizman D., Collins F.S., Liu P.P.,
RA   Chen X., Shohat M., Hamon M., Kahan T., Cercek A., Rotter J.I.,
RA   Fischel-Ghodsian N., Richards N., Shelton D.A., Gumucio D.,
RA   Yokoyama Y., Mangelsdorf M., Orsborn A., Richards R.I., Ricke D.O.,
RA   Buckingham J.M., Moyzis R.K., Deaven L.L., Doggett N.A.;
RT   "Ancient missense mutations in a new member of the RoRet gene family
RT   are likely to cause familial Mediterranean fever.";
RL   Cell 90:797-807(1997).
RN   [2]
RP   SEQUENCE OF 305-754 FROM N.A., AND VARIANTS FMF.
RX   MEDLINE=97434208; PubMed=9288094;
RA   Bernot A., Clepet C., Dasilva C., Devaud C., Petit J.-L.,
RA   Caloustian C., Cruaud C., Samson D., Pulcini F., Weissenbach J.,
RA   Heilig R., Notanicola C., Domingo C., Rozenbaum M., Benchetrit E.,
RA   Topaloglu R., Dewalle M., Dross C., Hadjari P., Dupont M.,
RA   Demaille J., Touitou I., Smaoui N., Nedelec B., Mery J.-P.,
RA   Chaabouni H., Delpech M., Grateau G.;
RT   "A candidate gene for familial Mediterranean fever.";
RL   Nat. Genet. 17:25-31(1997).
RN   [3]
RP   VARIANTS FMF, AND VARIANT GLN-202.
RX   MEDLINE=98334552; PubMed=9661175;
RA   Bernot A., da Silva C., Petit J.-L., Cruaud C., Caloustian C.,
RA   Castet V., Ahmed-Arab M., Dross C., Dupont M., Cattan D., Smaoui N.,
RA   Dode C., Pecheux C., Nedelec B., Medaxian J., Rozenbaum M., Rosner I.,
RA   Delpech M., Grateau G., Demaille J., Weissenbach J., Touitou I.;
RT   "Non-founder mutations in the MEFV gene establish this gene as the
RT   cause of familial Mediterranean fever (FMF).";
RL   Hum. Mol. Genet. 7:1317-1325(1998).
RN   [4]
RP   VARIANTS FMF ILE-680; ILE-681; ILE-694; VAL-694; MET-694 DEL AND
RP   ALA-726.
RX   MEDLINE=99149053; PubMed=10024914;
RA   Booth D.R., Gillmore J.D., Booth S.E., Pepys M.B., Hawkins P.N.;
RT   "Pyrin/marenostrin mutations in familial Mediterranean fever.";
RL   QJM 91:603-606(1998).
RN   [5]
RP   VARIANTS FMF.
RX   MEDLINE=99192341; PubMed=10090880;
RA   Aksentijevich I., Torosyan Y., Samuels J., Centola M., Pras E.,
RA   Chae J.J., Oddoux C., Wood G., Azzaro M.P., Palumbo G., Giustolisi R.,
RA   Pras M., Ostrer H., Kastner D.L.;
RT   "Mutation and haplotype studies of Familial Mediterranean Fever reveal
RT   new ancestral relationships and evidence for a high carrier frequency
RT   with reduced penetrance in the Ashkenazi Jewish population.";
RA   Am. J. Hum. Genet. 64:949-962(1999).
RN   [6]
RP   VARIANTS FMF GLN-148; SER-369; GLN-408; LEU-479; ILE-680; VAL-694;
RP   ALA-726 AND HIS-761.
RX   MEDLINE=99294585; PubMed=10364520;
RA   Cazeneuve C., Sarkisian T., Pecheux C., Dervichian M., Nedelec B.,
RA   Reinert P., Ayvazyan A., Kouyoumdjian J.-C., Ajrapetyan H.,
RA   Delpech M., Goossens M., Dode C., Grateau G., Amselem S.;
RT   "MEFV-Gene analysis in Armenian patients with Familial Mediterranean
RT   fever: diagnostic value and unfavorable renal prognosis of the M694V
RT   homozygous genotype-genetic and therapeutic implications.";
RL   Am. J. Hum. Genet. 65:88-97(1999).
RN   [7]
RP   VARIANTS FMF ILE-680; ILE-694; VAL-694 AND ALA-726.
RX   MEDLINE=99250763; PubMed=10234504;
RA   Shohat M., Magal N., Shohat T., Chen X., Dagan T., Mimouni A.,
RA   Danon Y., Lotan R., Ogur G., Sirin A., Schlezinger M., Halpern G.J.,
RA   Schwabe A., Kastner D., Rotter J.I., Fischel-Ghodsian N.;
RT   "Phenotype-genotype correlation in familial Mediterranean fever:
RT   evidence for an association between Met694Val and amyloidosis.";
RL   Eur. J. Hum. Genet. 7:287-292(1999).
RN   [8]
RP   VARIANTS FMF GLN-148; ILE-680; ILE-694; VAL-694; ARG-695; ALA-726 AND
RP   HIS-761.
RX   MEDLINE=20081069; PubMed=10612841;
RA   Akar N., Misiroglu M., Yalcinkaya F., Akar E., Cakar N., Tuner N.,
RA   Akcakus M., Taştan H., Matzner Y.;
RT   "MEFV mutations in Turkish patients suffering from Familial
RT   Mediterranean Fever.";
RL   Hum. Mutat. 15:118-119(2000).
RN   [9]
RP   VARIANT GLN-148.
```

```
RX   MEDLINE=20202844; PubMed=10737995;
RA   Ben-Chetrit E., Lerer I., Malamud E., Domingo C., Abeliovich D.;
RT   "The E148Q mutation in the MEFV gene: is it a disease-causing mutation
RT   or a sequence variant?";
RL   Hum. Mutat. 15:385-386(2000).
CC   -!- FUNCTION: LIKELY CONTROLS THE INFLAMMATORY RESPONSE IN
CC       DIFFERENTIATED GRANULOCYTES.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PERIPHERAL BLOOD
CC       LEUCOCYTES, PARTICULARLY IN MATURE GRANULOCYTES BUT NOT IN
CC       MONOCYTES AND LYMPHOCYTES. NOT EXPRESSED IN SPLEEN, THYMUS,
CC       PROSTATE, TESTIS, OVARY, SMALL INTESTINE, COLON, HEART, BRAIN,
CC       PLACENTA, LUNG, LIVER, MUSCLE, KIDNEY, PANCREAS.
CC   -!- DISEASE: Defects in MEFV are the cause of familial Mediterranean
CC       fever (FMF) [MIM:249100], an autosomal recessive inherited
CC       disorder characterized by recurrent episodic fever, serosal
CC       inflammation and pain in the abdomen, chest or joints. It is
CC       frequently complicated by amyloidosis, which leads to renal
CC       failure and can be prophylactically treated with colchicine. FMF
CC       primarily affects ethnic groups living around the Mediterranean
CC       basin: North-African Jews, Armenians, Arabs and Turks.
CC   -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- DATABASE: NAME=GeneDis; NOTE=FMF;
CC       WWW="http://life2.tau.ac.il/GeneDis/Tables/FMF/fmf.html".
CC   -----------------------------------------------------------------------
```

```
DR   EMBL; AF018080; AAB70557.1; -.
DR   EMBL; Y14441; CAA74793.1; -.
DR   EMBL; AJ003147; CAA05906.1; -.
DR   EMBL; AF111163; AAD26152.1; -.
DR   Genew; HGNC:6998; MEFV.
     MIM; 249100;
DR   GO; GO:0005875; C:microtubule associated complex; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0003779; F:actin binding activity; IDA.
DR   GO; GO:0008270; F:zinc ion binding activity; NAS.
DR   GO; GO:0006954; P:inflammatory response; IDA.
DR   InterPro; IPR001870; B302.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   InterPro; IPR006574; PRY.
DR   InterPro; IPR003877; SPRY receptor.
DR   InterPro; IPR000315; Znf_Bbox.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   Pfam; PF00622; SPRY; 1.
DR   Pfam; PF00643; zf-B_box; 1.
DR   SMART; SM00336; BBOX; 1.
DR   SMART; SM00589; PRY; 1.
DR   SMART; SM00449; SPRY; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
KW   Nuclear protein; Zinc-finger; Polymorphism; Disease mutation.
FT   DOMAIN        1     92       DAPIN.
FT   ZN_FING     370    412       B BOX-TYPE.
FT   DOMAIN      420    437       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   VARIANT     148    148       E -> Q (in FMF; common mutation;
FT                                associated with Ser-369 and Gln-408 in
FT                                cis; could be a polymorphism)
FT                                /FTId=VAR_009051.
FT   VARIANT     167    167       E -> D (in FMF).
FT                                /FTId=VAR_009052.
FT   VARIANT     202    202       R -> Q (in dbSNP:2242222).
FT                                /FTId=VAR_009053.
FT   VARIANT     267    267       T -> I (in FMF).
FT                                /FTId=VAR_009054.
FT   VARIANT     369    369       P -> S (in FMF; reduced penetrance among
FT                                Ashkenazi Jews; associated with Gln-148
FT                                and Gln-408 in cis; could be a
FT                                polymorphism).
FT                                /FTId=VAR_009055.
FT   VARIANT     408    408       R -> Q (in FMF; associated with Gln-148
FT                                and Ser-369 in cis; could be a
FT                                polymorphism).
FT                                /FTId=VAR_009056.
FT   VARIANT     479    479       F -> L (in FMF).
FT                                /FTId=VAR_009057.
FT   VARIANT     680    680       M -> I (in FMF).
FT                                /FTId=VAR_009058.
FT   VARIANT     681    681       T -> I (in FMF).
FT                                /FTId=VAR_009059.
FT   VARIANT     692    692       Missing (in FMF).
FT                                /FTId=VAR_009060.
FT   VARIANT     694    694       M -> I (in FMF).
FT                                /FTId=VAR_009061.
FT   VARIANT     694    694       M -> V (in FMF; very common mutation
FT                                particularly in North-African Jews; can
FT                                be associated with amyloidosis
FT                                development).
FT                                /FTId=VAR_009062.
FT   VARIANT     695    695       K -> R (in FMF; reduced penetrance among
FT                                Ashkenazi Jews).
FT                                /FTId=VAR_009063.
FT   VARIANT     695    695       V -> A (in FMF; common mutation; found in
FT                                Iraqi and Ashkenazi Jews, Druze,
FT                                Armenians).
FT                                /FTId=VAR_009064.
FT   VARIANT     726    726       M -> I (in FMF).
FT                                /FTId=VAR_009065.
```

```
RESULT 9
PYA3_HUMAN
ID   PYA3_HUMAN    STANDARD;    PRT;    980 AA.
AC   Q8WX94;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   PYRIN-containing APAF1-like protein 3.
GN   PYPAF3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22162427; PubMed=12019269;
RA   Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA   Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT   "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT   activation of NF-kappa B and caspase-1-dependent cytokine
RT   processing.";
RL   J. Biol. Chem. 277:29874-29880(2002).
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC   -----------------------------------------------------------------------
```

```
Query Match      12.4%;  Score 122.5;  DB 1;  Length 781;
Best Local Similarity   30.8%;  Pred. No. 0.0055;
Matches   33;  Conservative   18;  Mismatches   53;  Indels   3;  Gaps   1;

Qy      6  DAILDALENLTAEELKKFKLKLLSVPLRBGYGRIPRGALLSMDALDLCIDKLVSFYLETYG   65
           :||  |||| | : |:||  |  |:||  | ||  : ||  | :|| |  ||::| | :|
Db      7  DHLLSTLEELVPYDFEKFKFKFKLQNTSVQXEHSRIPRSQIQRARPVKMATLLVTTYGEEYA   66

Qy     66  AEITANVLRDMGLGEMAGQLQAATHQGSGAAPAGIQAPPQSAAXPGL   112
           :||  |||| |:  ||  : |:|:||  | ||  |||
Db     67  VQLTLQVLRAINQRLIAEELHRAAIQEYSTQENGTD---DSAASSSL   110
```

CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF464765; AAL69963.1; -.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   ATP-binding; Leucine-rich repeat; Repeat.
FT   DOMAIN   1   93   DAPIN.
FT   DOMAIN   172   491   NACHT.
FT   REPEAT   614   638   LRR 1.
FT   REPEAT   674   697   LRR 2.
FT   REPEAT   760   784   LRR 3.
FT   REPEAT   788   810   LRR 4.
FT   REPEAT   817   840   LRR 5.
FT   REPEAT   845   868   LRR 6.
FT   REPEAT   874   897   LRR 7.
FT   REPEAT   902   928   LRR 8.
FT   REPEAT   933   957   LRR 9.
FT   REPEAT   957   957   LRR 8.
FT   NP_BIND   178   185   ATP (POTENTIAL).
SQ   SEQUENCE   980 AA; 111806 MW; 822A2FD4338003D CRC64;

Query Match   10.1%; Score 99.5; DB 1; Length 980;
Best Local Similarity   26.6%; Pred. No. 0.71;
Matches   45; Conservative   21; Mismatches   80; Indels   23; Gaps   5;

Qy   12 LENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYGAELTAN 71
               || ||::|:|: : | ::| ||::: |: : || : |:: :
Db   14 LEQLNEDELKSFKSLLWAPPLEDVLQKTPWSEVEEADGKKLAEILVNTSSENWIRNATVN 73

Qy   72 VLRDMGLQEMA--------------GGLQAATHQGSGAAPAGIQAPFQSAAKPGLH---FIDCH 118
               :|:: ||::| || | : | :| |:: ||:||
Db   74 ILEEMNLTELCKMAKAEMMEDGQVQEIDNPELGDA-----EEDSELAKPGEKGGWRNSME 128

Qy   119 RAALIARVTNVEWLLDALYGKVLITDEQYQAVRABP-TNPSKMRKLFSFT 166
               ::|:|: :: | ||| || : :| : |: :| :|| |:||:|
Db   119 RAALIARVTNVEWLLDALYGKVLITDEQYQAVRABP-TNPSKMRKLFSFT 166

Qy   129 KQSLVWKNTFWQGDIDNFHDDVTLRNQ----RFIPFLNPRTPRKLTPYT 173

RESULT 10

ID   CAR8_HUMAN   STANDARD;   PRT;   431 AA.
AC   Q9Y2G2; Q96P82;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Caspase recruitment domain protein 8 (Apoptotic protein NDPP1) (DACAR)
DE   (CARD-inhibitor of NF-kappaB activating ligand) (CARDINAL) (TUCAN).
GN   CARD8 OR NDPP1 OR KIAA0955.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=21950691; PubMed=11821383;

RA   Razmara M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
RA   Distefano P.S., Bertin J., Alnemri E.S.;
RT   "CARD-8 protein, a new CARD family member that regulates caspase-1
RT   activation and apoptosis.";
RL   J. Biol. Chem. 277:13952-13958(2002).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RA   Zhang H.;
RT   "A novel apoptotic protein, NDPP1, containing CARD and BH3 domains.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RA   Guiet C., Vito P.;
RT   "DACAR, a novel CARD-containing protein.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=21570185; PubMed=11551959;
RA   Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
RA   MacFarlane M., Martin S.J.;
RT   "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
RT   of multiple caspase recruitment pathways.";
RL   J. Biol. Chem. 276:44069-44077(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Kidney;
RA   Guo J.H., Yu L.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   CHARACTERIZATION.
RX   MEDLINE=21402909; PubMed=11408476;
RA   Pathan N., Marusawa H., Krajewska M., Matsuzawa S.-I., Kim H.,
RA   Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Pio F.,
RA   Godzik A., Reed J.C.;
RT   "TUCAN, an antiapoptotic caspase-associated recruitment domain family
RT   protein overexpressed in cancer.";
RL   J. Biol. Chem. 276:32220-32229(2001).
RN   [8]
RP   CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
RX   MEDLINE=22062958; PubMed=12067710;
RA   Stilo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
RT   "TUCAN/CARDINAL and DRAL participate in a common pathway for
RT   modulation of NF-kappaB activation.";
RL   FEBS Lett. 521:165-169(2002).
CC   -!- FUNCTION: Inhibits NF-kappaB activation. May participate in a
CC   regulatory mechanism that coordinates cellular responses
CC   controlled by NF-kappaB transcription factor. Involved in the
CC   negative regulation of caspase-1.
CC   -!- SUBUNIT: May form homodimers. Interacts with NEMO and DRAL. Binds
CC   to caspase-1, pseudo-ICE and ICEBERG.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=Long;
CC   IsoId=Q9Y2G2-1; Sequence=Displayed;
CC   Name=Short;
CC   IsoId=Q9Y2G2-2; Sequence=VSP_000782, VSP_000783;
CC   -!- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
CC   placenta. Lower expression in heart, kidney and liver. Also
CC   expressed in spleen, lymph node and bone marrow.
CC   -!- SIMILARITY: Contains 1 CARD domain.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AB023172; BAA76799.1; -.
DR   EMBL; AF322104; AAG50014.1; -.
DR   EMBL; AF331519; AAK01126.1; -.

```
DR   EMBL; AY026322; AAK08982.1; -.
DR   EMBL; AF405558; AAL02427.1; -.
DR   EMBL; AF511652; AAM46959.1; -.
DR   InterPro; IPR001315; CARD.
DR   Pfam; PF00619; CARD; 1.
DR   SMART; SM00114; CARD; 1.
DR   PROSITE; PS50209; CARD; 1.
KW   Apoptosis; Nuclear protein; Alternative splicing.
FT   DOMAIN       340    430       CARD.
FT   VARSPLIC     282    286       ELKLS -> WISSL (in isoform Short).
FT                                 /FTId=VSP_000782.
FT   VARSPLIC     287    431       Missing (in isoform Short).
FT                                 /FTId=VSP_000783.
FT   MUTAGEN      366    366       L->R: INHIBITS HOMODIMER FORMATION.
FT   CONFLICT      60     60       E -> G (IN REF. 5).
FT   CONFLICT     326    326       V -> M (IN REF. 5).
FT   CONFLICT     422    422       L -> P (IN REF. 5).
SQ   SEQUENCE    431 AA;  48932 MW;  CB54D13080773286 CRC64;

Query Match             10.0%;  Score 99;  DB 1;  Length 431;
Best Local Similarity   24.0%;  Pred. No. 0.3;
Matches   42;  Conservative   32;  Mismatches   73;  Indels   28;  Gaps    5;

Qy     21 KKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQE 80
          |||||||||||||||||||||||||||||
Db    281 KELKLSYRSPGEIQHFSKFYAGQMKEPIQLEITEK-------RHGTLVWDTEVKPVDLQL 333

Qy     81 MAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALIARVTNVEWLLDALY-GK 139
          :|  |:  .:|:   |:|||::||:    :||
Db    334 VA--------ASAP---PPFS---GAAFVKENHRQIQARMGDLKGVLDDLQDNE 373

Qy    140 VLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQALRESQSYLVEDLER 194
          |||:|::|||:::.|:|:  |:|: :::|::: |||::||:||
Db    374 VLTENEKELVEQEKTRQSKNEALLSMVEKKGDLALDVLFRSISERDPYLVSYLRQ 428

RESULT 11
ID   SILE_SALTY     STANDARD;      PRT;   143 AA.
AC   Q9Z4N3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Silver-binding protein silE precursor.
GN   SILE.
OS   Salmonella typhimurium.
OG   Plasmid pMG101.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 21-35, AND CHARACTERIZATION.
RX   MEDLINE=99128056; PubMed=9930866;
RA   Gupta A., Matsui K., Lo J.-F., Silver S.;
RT   "Molecular basis for resistance to silver cations in Salmonella.";
RL   Nat. Med. 5:183-188(1999).
-!-  FUNCTION: COMPONENT OF THE SIL CATION-EFFLUX SYSTEM THAT CONFERS
-!-  RESISTANCE TO SILVER.
-!-  SUBCELLULAR LOCATION: Periplasmic.
-!-  INDUCTION: By silver.
-!-  SIMILARITY: TO E.COLI PCOE.
```

```
DR   EMBL; AF067954; AAD11743.1; -.
KW   Metal-binding; Periplasmic; Signal; Plasmid.
FT   SIGNAL        1     20
FT   CHAIN        21    143       SILVER-BINDING PROTEIN SILE.
SQ   SEQUENCE    143 AA;  15201 MW;  703B9BC0F8BDDCBE CRC64;

Query Match              9.2%;  Score 91.5;  DB 1;  Length 143;
Best Local Similarity   31.6%;  Pred. No. 0.37;
Matches   31;  Conservative   11;  Mismatches   47;  Indels    9;  Gaps    4;

Qy     64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIDQHRAAL 122
          :|  |:|  |:  :|||:  ::   |||:  :|:|
Db     12 FGLISSAWATETVNIHERVNNAQAPAHQMQSAAPVGIQ-----GTAPRMAGMQHEQAI 66

Qy    123 IARVTNVEWLLDALYGKVLTDEQYQ--AVRAEPTNPSK 158
          |:||  ||  |:  ::|||
Db     67 IAHETWTNGSADA-HQKMVESHQRMMGSQTVSFTGPSK 103

RESULT 12
ID   RUVB_THETN     STANDARD;      PRT;   338 AA.
AC   Q8RAN2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Holliday junction DNA helicase ruvB.
GN   RUVB OR TTE1180.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC   Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STRAIN=MB4 / JCM 11007;
RX   MEDLINE=21998816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
-!-  FUNCTION: The ruvA-ruvB complex in the presence of ATP renatures
     cruciform structure in supercoiled DNA with palindromic sequence,
     indicating that it may promote strand exchange reactions in
     homologous recombination. RuvAB is a helicase that mediates the
     Holliday junction migration by localized denaturation and
     reannealing (By similarity).
-!-  SUBUNIT: Forms a complex with ruvA (By similarity).
-!-  SIMILARITY: BELONGS TO THE RUVB FAMILY.
```

```
DR   EMBL; AE013080; AAM24411.1; -.
DR   HAMAP; MF_00016; -; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003959; AAA_ATPase_centr.
DR   InterPro; IPR004605; RuvB.
DR   Pfam; PF00004; AAA; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR00635; ruvB; 1.
KW   DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW   Complete proteome.
FT   NP_BIND      59     66       ATP (POTENTIAL).
SQ   SEQUENCE    338 AA;  38072 MW;  2B9562D8CBD988D1 CRC64;

Query Match              9.0%;  Score 89;  DB 1;  Length 338;
Best Local Similarity   24.1%;  Pred. No. 1.7;
Matches   39;  Conservative   32;  Mismatches   73;  Indels   18;  Gaps    5;

Qy      3 RARDAILDALENLTREELK---KFKLKLLSVPLREG---YGRIPRGALLSMDAL----- 50
```

```
Db   169 RSRFGVINRLDYYSVEELKEIIKRSANILNIGIDEDAAFEIARRSRGTPRIANRLLKRVR 228
          |:|  ::|  | |||| |:   :| :||
Qy   51  DLTDKLVSFYLETYGAELTANVL--RDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAA 108
          |:|  ::|  |
Db   229 DFAEVKGNGYIDYNTANIALNMLGVDEMGLEEIDRKILIAIIEKFGGGPVGIDAIAASVG 288
          :||   |:
Qy   109 KPGLHFIDQHRAALIARVTNVEWLLLDALIYGKVLTDBQYQAVR 150
          ||  :   |
Db   289 EDGDTIEDMYEPYLM----QIGFLNRTPRGRVVTKLAYQYLK 326
```

RESULT 13

ID   M22_STRPY   STANDARD;   PRT;   372 AA.
AC   P50469;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   M protein, serotype 2.2 precursor.
GN   EMM2.2.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fischetti V.A., Bessen D.E.;
RT   "Immunoglobulin A binding protein.";
RL   Patent number US5556944, 17-SEP-1996.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bessen D.E., Fischetti V.A.;
RX   MEDLINE=92104662; PubMed=1370269;
RA   Bessen D.E., Fischetti V.A.;
RT   "Nucleotide sequences of two adjacent M or M-like protein genes of
RT   group A streptococci: different RNA transcript levels and
RT   identification of a unique immunoglobulin A-binding protein.";
RL   Infect. Immun. 60:124-135(1992).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC   OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC   THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC   PHAGOCYTOSIS.
CC   -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC   an amide bond (Potential).
CC   -!- SIMILARITY: TO OTHER M PROTEINS.

DR   EMBL; X61276; CAA43582.1; -.
DR   EMBL; I26204; -; NOT_ANNOTATED_CDS.
DR   PIR; S23326; S23326.
DR   InterPro; IPR005877; Gpos_YSIRK.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR006192; LPXTG.
DR   InterPro; IPR003345; M_repeat.
DR   Pfam; PF00746; Gram_pos_anchor; 1.
DR   Pfam; PF02370; M; 8.
DR   PRINTS; PR00015; GPOSANCHOR.
DR   Pfam; PF04650; YSIRK_signal; 1.
DR   TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR   TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR   PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW   Virulence; Phagocytosis; Cell wall; Peptidoglycan-anchor; Repeat;
KW   Antigen; Coiled coil; Signal.
FT   SIGNAL   1   41   POTENTIAL.
FT   CHAIN   42   342   M PROTEIN, SEROTYPE 2.2.
FT   PROPEP   343   372   REMOVED BY SORTASE (POTENTIAL).
FT   DOMAIN   131   244   3 X REPEATS, TYPE C.
FT   REPEAT   131   153   C-1.
FT   REPEAT   173   195   C-2.
FT   REPEAT   222   244   C-3.
FT   DOMAIN   305   338   GLY/PRO-RICH (CELL WALL-SPANNING)
FT   (BY SIMILARITY)
FT   SITE   339   343   LPXTG SORTING SIGNAL (POTENTIAL).
FT   MOD_RES   342   342   AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ   SEQUENCE   372 AA;   41149 MW;   E8FD5D0920C95C74 CRC64;

Query Match   53;   Score 86.5;   DB 1;   Length 372;
Best Local Similarity   26.8%;   Pred. No. 3.1;
Matches   53;   Conservative   30;   Mismatches   78;   Indels   37;   Gaps   9;

```
Qy   12  LENLTAEELKKFKLLSVPLRFGYGRIPRGALLSMDAL-----DLTDKLVSFYLETYGA 66
          |||   :|| ::|  : :           :|:|  |            :|| |
Db   78  LEKINAEEEKNKLEAINKELNENYYKLQDG----IDALEKEKEDLKTTLAKITKENIS 133

Qy   67  ELTANVL-RDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQ--SAAKPGL-HFIDQHRAAL 122
          :|   |:     :|      :    |:   |              :|| ||
Db   134 EASRKGLSRDLEASRTAKKELEAKHQKLEAENKKLTEGNQVSEASRKGLSNDLEASRAA- 192

Qy   123 IARVTNVEWLLLDALIYGKVLTDEQ------YQAVRAEPTNPSKMRKLFSFTPAWWTCKDLL 177
          ||: |:|| |||||||||||:||:        :
Db   193 ----KKELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLS-----------RD-- 234

Qy   178 LQALRESQSYLVEDLERS 195
          | |:| |:  |: |
Db   235 LEASREANKKVTSELTQA 252
```

RESULT 14

ID   SYFB_CORGL   STANDARD;   PRT;   835 AA.
AC   Q8NQN6;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE   (Phenylalanine--tRNA ligase beta chain) (PheRS).
GN   PHET OR CGL1390.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC   diphosphate + L-phenylalanyl-tRNA(Phe).
CC   -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC   -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC   similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
CC   family. Subfamily 1.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.

DR   EMBL; AP005278; BAB98783.1; ALT_INIT.
DR   HAMAP; MF_00283; -; 1.
DR   InterPro; IPR005146; B3_4.

```
DR   InterPro; IPR005147; B5.
DR   InterPro; IPR005121; Fdx-AntiCB.
DR   InterPro; IPR004532; PheT_bact.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF03483; B3_4; 1.
DR   Pfam; PF03484; B5; 1.
DR   Pfam; PF03147; FDX-ACB; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   TIGRFAMs; TIGR00472; pheT_bact; 2.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW   Metal-binding; Magnesium; RNA-binding; tRNA-binding;
KW   Complete proteome.
FT   DOMAIN       44    160       TRNA-BINDING.
FT   METAL       472    472       MAGNESIUM (BY SIMILARITY).
FT   METAL       478    478       MAGNESIUM (VIA CARBONYL OXYGEN) (BY
FT                                SIMILARITY).
FT   METAL       481    481       MAGNESIUM (BY SIMILARITY).
FT   METAL       482    482       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   835 AA;  89435 MW;  285EC9A0673DA49F CRC64;

Query Match             8.7%;  Score 86.5;  DB 1;  Length 835;
Best Local Similarity   24.5%;
Matches   40;  Conservative   29;  Mismatches   73;  Indels   21;  Gaps    7;

Qy     29 SVPLREGYGRIPRGALLSMDALDLTDKLVSFYL---ETYGAELTANVLRDMGLQEMAQOL 85
Db    114 AISARETYGRMSAGMICSASELGLADKQNSGITTLDPSYG-EPGBDARQALGLBDTVFDV 172

Qy     86 QAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALIARVTNVBWLLDALYGKVLT--- 142
Db    173 NVTPDRGYALSARGLTRELASAF--SLIFTD--PAIEPAVAGIEVKVPAVEGSLINVEL 227

Qy    143 DEQYQAVRAEPTNPSKMRKLFSPTPAWN----WTCKDLLLQALR 182
Db    228 REETKAIR------FGLRKVSGIDPAAESPFWMQRELMLSGQR 264
```

```
RESULT 15
CIS1_MOUSE
ID   CIS1_MOUSE     STANDARD;     PRT;  1033 AA.
AC   Q8R4B8;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing
DE   APAF1-like protein 1) (Mast cell maturation inducible protein 1).
GN   CIAS1 OR PYPAF1 OR MMIG1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/CJ;
RA   Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA   Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA   Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT   "Identification of inducible genes during in vitro maturation of mouse
RT   bone marrow-derived mast cells to connective tissue-type mast cells.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: May function as a potential inducer of apoptosis.
CC       Interacts selectively with apoptosis-associated specklike protein
CC       containing a CARD domain (ASC). This complex may function as an
CC       upstream activator of NF-kappaB signaling (By similarity).
CC   -!- SIMILARITY: Contains 1 DAPIN domain.
CC   -!- SIMILARITY: Contains 1 NACHT domain.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF486632; AAL90874.1; -.
DR   MGD; MGI:2653833; Cias1.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007091; LRR_RNinh.
DR   InterPro; IPR007091; LRR_RNinh_sub.
DR   InterPro; IPR003590; LRR_RNinh_sub.
DR   InterPro; IPR007111; NACHT_NTPase.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00368; LRR_RI; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; Repeat; Leucine-rich repeat.
FT   DOMAIN        1     91       DAPIN.
FT   DOMAIN      216    532       NACHT.
FT   REPEAT      737    760       LRR 1.
FT   REPEAT      794    817       LRR 2.
FT   REPEAT      851    874       LRR 3.
FT   REPEAT      880    903       LRR 4.
FT   REPEAT      908    931       LRR 5.
FT   REPEAT      937    964       LRR 6.
FT   REPEAT      965    988       LRR 7.
SQ   SEQUENCE   1033 AA;  118274 MW;  5924690966B12117 CRC64;

Query Match             8.7%;  Score 86;  DB 1;  Length 1033;
Best Local Similarity   35.6%;
Matches   21;  Conservative    9;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSF 59
Db      1 MTSVRCKLAQYLEDLEDVDLKKFKMHLEDYPPEKGCIPVPRGQMEKADHLDLATLMIDF 59
```

Search completed: January 29, 2004, 13:46:32
Job time : 5.00185 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.   GenCore version 5.1.6

Run on: January 29, 2004, 13:06:01 ; Search time 10.8067 seconds
                                            (without alignments)
                                            4656.416 Million cell updates/sec

Title: US-09-996-617-8
Perfect Score: 990
Sequence: 1 MGRARDAILDALENLTAEEL..........LLLQALRESQSYLVEDLERS 195
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_23:*
                1: sp_archea:*
                2: sp_bacteria:*
                3: sp_fungi:*
                4: sp_human:*
                5: sp_invertebrate:*
                6: sp_mammal:*
                7: sp_mhc:*
                8: sp_organelle:*
                9: sp_phage:*
                10: sp_plant:*
                11: sp_rodent:*
                12: sp_virus:*
                13: sp_vertebrate:*
                14: sp_unclassified:*
                15: sp_rvirus:*
                16: sp_bacteria:*
                16: sp_bacteriap:*
                17: sp_archeap:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	% Length	DB	ID	Description
1	736	74.3	195	6	Q8HXK9	Q8hxk9 bos taurus
2	673	68.0	193	11	Q8CHK8	Q8chk8 rattus norv
3	275	27.8	89	4	Q8NFP8	Q8nfp8 homo sapien
4	267	27.0	89	4	Q8WXC3	Q8wxc3 homo sapien
5	142	14.3	750	11	Q9JJ25	Q9jj25 rattus norv
6	142	14.3	767	11	Q9JZ26	Q9jz26 mus musculu
7	104	10.5	404	13	Q9DDJ2	Q9ddj2 brachydanio
8	91.5	9.2	127	2	Q8VPY3	Q8vpy3 serratia ma
9	91.5	9.2	127	2	Q8VPW8	Q8vpw8 salmonella
10	91.5	9.2	127	2	Q8VPX8	Q8vpx8 serratia ma
11	91.5	9.2	127	2	Q8VPY8	Q8vpy8 serratia ma
12	91.5	9.2	383	13	Q919L7	Q919l7 brachydanio
13	86.5	8.7	439	16	P74033	P74033 synechocyst
14	86.5	8.7	828	16	Q8NQN6	Q8nqn6 corynebacte
15	86	8.7	580	17	Q8TSR7	Q8tsr7 methanosarc
16	86	8.7	1033	11	Q8R4B8	Q8r4b8 mus musculu

ALIGNMENTS

RESULT 1
ID   Q8HXK9
AC   Q8HXK9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Apoptosis-associated speck-like protein containing a CARD.
GN   BASC.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Masumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,
RA   Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazelett P.,
RA   Postlethwalt J.H., Nunez G., Inohara N.;
RT   "Caspy: A Zebrafish caspase activated by ASC oligomerization required
RT   for pharyngeal Arch development.";
RL   J. Biol. Chem. 0:0-0(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Masumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T.,
RA   Nikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.;
RT   "ASC, a novel 22-kDa protein, aggregates during apoptosis of human
RT   promyelocytic leukemia HL-60 cells.";
RL   J. Biol. Chem. 274:33835-33838(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;
RX   MEDLINE=20580347; PubMed=11139337;
RT   "Murine ortholog of ASC, a CARD-containing protein, self-associates,
RT   and exhibits restricted distribution in developing mouse embryos.";
RL   Exp. Cell Res. 262:128-133(2001).
DR   EMBL; AB050006; BAC43753.1; -.
SQ   SEQUENCE 195 AA; 21917 MW; 7C9D4BD8DBA9A9E8 CRC64;

PRELIMINARY;   PRT;   195 AA.

17	85	8.6	1488	16	Q8Z7Z5	Q8z7z5 salmonella
18	85	8.6	1488	16	Q93S87	Q93s87 salmonella
19	84.5	8.5	894	10	Q8GSF4	Q8gsf4 oryza sativ
20	84.5	8.5	1241	3	Q9HEM2	Q9hem2 neurospora
21	84	8.5	292	16	O25976	O25976 helicobacte
22	84	8.5	530	5	Q8T9T6	Q8t9t6 aedes aegyp
23	83.5	8.4	542	17	Q9HRG1	Q9hrg1 halobacteri
24	83.5	8.4	2556	12	Q9QH56	Q9qh56 gallid herp
25	83	8.4	469	16	Q8CVN9	Q8cvn9 escherichia
26	83	8.4	545	16	Q9CBL2	Q9cbl2 mycobacteri
27	83	8.4	877	2	Q45095	Q45095 bacillus ci
28	82.5	8.3	409	5	O96984	O96984 stylonychia
29	82.5	8.3	522	16	O85210	O85210 vibrio chol
30	82.5	8.3	522	16	Q9K394	Q9k394 vibrio chol
31	82.5	8.3	567	16	Q8NIK9	Q8nik9 cryptococcu
32	82	8.3	292	16	Q9ZJI1	Q9zji1 helicobacte
33	81	8.2	456	4	Q9BUC6	Q9buc6 homo sapien
34	81	8.2	503	2	Q9XBW8	Q9xbw8 rhodobacter
35	81	8.2	631	4	Q8NDM4	Q8ndm4 homo sapien
36	81	8.2	631	4	Q96920	Q96920 homo sapien
37	81	8.2	792	16	Q82PS3	Q82ps3 salmonella
38	81	8.2	792	16	Q8Z6J0	Q8z6j0 salmonella
39	80.5	8.2	896	2	Q9AN79	Q9an79 bradyrhizob
40	80.5	8.2	9477	2	Q9L4X3	Q9l4x3 streptomyce
41	80.5	8.1	266	10	Q8LM81	Q8lm81 oryza sativ
42	80.5	8.1	718	2	Q9JYB6	Q9jyb6 neisseria m
43	80.5	8.1	725	16	Q9JT97	Q9jt97 neisseria m
44	80.5	8.1	727	16	Q8DE24	Q8de24 vibrio vuln
45	80.5	8.1	972	4	Q8N612	Q8n612 homo sapien

```
Qy                      74.3%;   Score 736; DB 6;   Length 195;
Query Match
Best Local Similarity   75.4%;   Pred. No. 1.7e-59;
Matches 147;  Conservative 20;  Mismatches 28;  Indels 0;  Gaps 0;

Qy    1 MGGTRDAILDALENLTADELKKFKMKLLSVPLREGYGRIPRGTLLPLDAVDLTDKLVSFY   60
          ||||||||||||||||||:|||||||||||||||||||||||||||||:||||||||||:|
Db    1 MGCTRDAILDALENLTADELKKFRMKLLSVPLREGYGRIPRGTLLPLDAVDLTDKLVSFY   60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA  120
          |||||||||||||||||:|||||||| |:  ||||||:|||||||||||||||||||||||
Db   61 LEAYGAELTALVLRDMGMQEVAEQLQETMSKGPRNVLAEVRDPLQKTAKPGLHPVDQHRA  120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAMWTCKDLLLQA   180
          ||||||| :||| ||||:|||::|||||||:|  ||||||||||||| ||| ||||:|
Db  121 ALIARVTVDGVLIDALYGKVLTEEQYQAVRAERTSSDKMRKLFSFSPAWNMTCKDLLLQA  180

Qy  181 LRESQSYLVEDLERS                                             195
          |:||||||||||||:|
Db  181 LRQTQPYLVTDLEQS                                             195


RESULT 2
ID   Q8CHK8          PRELIMINARY;    PRT;  193 AA.
AC   Q8CHK8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Apoptosis-associated speck-like protein.
GN   RASC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Masumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,
RA   Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazelett P.,
RA   Postlethwait J.H., Nunez G., Inohara N.;
RT   "Caspy: A Zebrafish caspase activated by ASC oligomerization required
RT   for pharyngeal Arch development.";
RL   J. Biol. Chem. 0:0-0(2002).
DR   EMBL; AB053165; BAC43754.1; -.
DR   PROSITE; PS50824; DAPIN; 1.
SQ   SEQUENCE   193 AA;  21654 MW;  F3B27B560D86A17B CRC64;

Query Match             68.0%;   Score 673; DB 11;  Length 193;
Best Local Similarity   69.7%;   Pred. No. 9.7e-54;
Matches 136;  Conservative 20;  Mismatches 37;  Indels 2;  Gaps 2;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY   60
          |||||||||||||||:|||||:||||||||||||||||||||||||:|||||||||||:|
Db    1 MGRARDAILDALENLTADEPKKFKMKLLTAPVREGYGRIPRGALLQMDPIDLTDKLVSYY   60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA  120
          ||:||:||| :||||||:||:|:|||   :: |||| : :  ||||| ||| |||:|
Db   61 LEGYGLELTMSVLRDMGIQELAEQLQKIMEE-SGAVATATSVPAQGTARTE-HFVDQHRQ  118

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAMWTCKDLLLQA   180
          ||||||| :| ||||||:|||: |||||||||  |:|:||||| :|| |||:||| ::
Db  119 ALIARVTKVDGLLDALYGNVLTEGQYQAVRAETPNQNMRKLPSPAPAWNLTCKNLFLRA  178

Qy  181 LRESQSYLVEDLERS                                             195
          |:|||||||||||:|
Db  179 LRQTQPYLVTDLEQS                                             193
```

```
Db   61 YEDYAAELVVAVLRDMRMLEEAARLQRA                                 88

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAA                                 88
          ||:|||||| :||||| :|:  :|||:|
Db    1 MGTKREAILKVLENLTPEELKKFKFKMKLGTVPLREGFGRIPRGALGQLDIVDLTDKLVSY   60

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY   60
          |||||||||||||||:|||||:|||||| ||||||:|||||||| |:| |:|||||||:|
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrin-domain containing protein.
GN   PYC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
```

```
RESULT 3
ID   Q8NFP8          PRELIMINARY;    PRT;  89 AA.
AC   Q8NFP8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
```

```
RESULT 4
ID   Q8WXC3          PRELIMINARY;    PRT;  89 AA.
AC   Q8WXC3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Pyrin-only protein 1.
GN   POP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bertin J.;
RT   "POP1 : a Pyrin-only protein that regulates inflammatory signaling.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF454669; AAL58439.1; -.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PROSITE; PS50824; PAAD_DAPIN; 1.
SQ   SEQUENCE   89 AA;  10107 MW;  4CDF6D6720DDD998E CRC64;

Query Match             27.0%;   Score 267; DB 4;   Length 89;
Best Local Similarity   64.8%;   Pred. No. 5.6e-17;
Matches 57;  Conservative 9;  Mismatches 22;  Indels 0;  Gaps 0;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY   60
          | :||| ||||:|||:|||||:||||||| ||||||:|||||||| ||| |:|||||||:|
Db    1 MGTKREAILKVLENLTPEELKKFKMKLGTVPLREGFGRIPRGALGQLDIVDLTDKLVSAY   60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAA                                 88
          ||:|||||| :||||| :|:  :|||:|
Db   61 YEDYAAELVVAVLRDMRMLEEAARLQRA                                 88
```

```
RESULT 5
ID   Q9JZ25          PRELIMINARY;    PRT;  750 AA.
AC   Q9JZ25;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrin (Marenostrin).
GN   MEFV.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=20279845; PubMed=10818206;
RA   Chae J.J., Centola M., Aksentijevich I., Dutra A., Tran M., Wood G.,
RA   Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
RT   "Isolation, genomic organization, and expression analysis of the mouse
RT   and rat homologs of MEFV, the gene for familial Mediterranean fever.";
RL   Mamm. Genome 11:428-435(2000).
CC   -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
CC       MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
CC       AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
CC       PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN AND, TO A LESSER DEGREE IN
CC       THE LUNG. NOT EXPRESSED IN THYMUS, TESTIS, OVARY, HEART, BRAIN,
CC       LIVER, KIDNEY AND MUSCLE.
CC   -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
DR   EMBL; AF143410; AAF03767.1; -.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00336; BBOX; 1.
DR   Pfam; PF00643; zf-B_box; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   InterPro; IPR000315; Znf_Bbox.
DR   MGD; MGI:1859396; Mefv.
KW   Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
KW   Microtubules; Zinc-finger; Zinc.
FT   DOMAIN        1     92       DAPIN.
FT   ZN_FING     440    482       B BOX-TYPE.
SQ   SEQUENCE   750 AA;  83994 MW;  C76D0F3E02711312 CRC64;
```

Query Match   14.3%; Score 142; DB 11; Length 750;
Best Local Similarity   33.7%; Pred. No. 0.0029;
Matches   35; Conservative   19; Mismatches   50; Indels   0; Gaps   0;

```
Qy    6 DAILDALENLTAEELKKFKLKILLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYG 65
          |:||:|.||||:|||.|.|::||:|| |:.|| |  ||:|.:|.::|  |::.|:.:|
Db    7 DHILNTLEELLPYELEKFKFKLHTTSLEKGHSRIPLSLVMARPIKLTRLLLITYYGEEYA 66

Qy   66 AELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAK 109
          |:|| ::|:|.|...| |:|         |: |:||:
Db   67 VRLTLQILRATNQRQLAEELHKATGPEHLTEENGVGGSVQSSAE 110
```

```
RA   Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
RT   "Isolation, genomic organization, and expression analysis of the mouse
RT   and rat homologs of MEFV, the gene for familial Mediterranean fever.";
RL   Mamm. Genome 11:428-435(2000).
CC   -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
CC       MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
CC       AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
CC       PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN PERIPHERAL BLOOD
CC       GRANULOCYTES. NOT EXPRESSED IN LYMPHOCYTES, THYMUS, TESTIS, OVARY,
CC       HEART, BRAIN, LUNG, LIVER, KIDNEY AND MUSCLE.
CC   -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC   -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
DR   EMBL; AF143409; AAF03766.1; -.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   SMART; SM00336; BBOX; 1.
DR   Pfam; PF00643; zf-B_box; 1.
DR   PROSITE; PS50824; DAPIN; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   InterPro; IPR000315; Znf_Bbox.
KW   Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
KW   Microtubules; Zinc-finger; Zinc.
FT   DOMAIN        1     92       DAPIN.
FT   ZN_FING     439    481       B BOX-TYPE.
SQ   SEQUENCE   767 AA;  86264 MW;  PF102CB3FPD7C1EB CRC64;
```

Query Match   14.3%; Score 142; DB 11; Length 767;
Best Local Similarity   28.3%; Pred. No. 0.0003;
Matches   32; Conservative   26; Mismatches   47; Indels   8; Gaps   1;

```
Qy    6 DAILDALENLTAEELKKFKLKILLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYG 65
          |:||:|.||||:|||.|.|::||:|| |:.|| |  ||:|.:|.::|  |::.|:.:|
Db    7 DHILNTLEELLPYDPEKFKFKLQNTSLEKGHSKIPRGHHMQMARPVKLASLLITYYGEEYA 66

Qy   66 AELTANVLRDMGLQEMAGQLQAAT----------HQGSGAAPAGIQAPPQSAAKD 110
          |:|| ::|:|.|...| |:|           |: |:||:
Db   67 VRLTLQILRATNQRQLAEELRKATGTEHLIEENRVGGSVQSSVENKAKSVKVP 119
```

```
RESULT 6
Q9JJ26
ID   Q9JJ26    PRELIMINARY;    PRT;   767 AA.
AC   Q9JJ26;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrin (Marenostrin).
GN   MEFV.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   STRAIN=129/Sv;
RX   MEDLINE=20279845; PubMed=10818206;
RA   Chae J.J., Centola M., Aksentijevich I., Dutra A., Tran M., Wood G.,
```

```
RESULT 7
Q9DDJ2
ID   Q9DDJ2    PRELIMINARY;    PRT;   404 AA.
AC   Q9DDJ2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Caspy2.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20373792; PubMed=10917738;
RA   Inohara N., Nunez G.;
RT   "Genes with homology to mammalian apoptosis regulators identified in
RT   zebrafish.";
RL   Cell Death Differ. 7:509-510(2000).
GN   CASPY2.
DE   CASPB OR CASPY2.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20373792; PubMed=10917738;
RA   Inohara N., Nunez G.;
DR   EMBL; AF327410; AAG45230.1; -.
DR   HSSP; P29466; 1ICE.
DR   ZFIN; ZDB-GENE-020812-1; caspb.
DR   InterPro; IPR002398; ICE.
DR   InterPro; IPR002138; ICE_P10.
DR   InterPro; IPR001309; ICE_P20.
DR   InterPro; IPR004020; PAAD_DAPIN_dom.
DR   Pfam; PF00656; ICE_P20; 1.
DR   Pfam; PF00655; ICE_P10; 1.
DR   Pfam; PF02758; PAAD_DAPIN_dom; 1.
```

```
DR   Pfam; PF02758; PAAD_DAPIN; 1.
DR   PRINTS; PR00376; IL1BCENZYME.
DR   SMART; SM00115; CASC; 1.
DR   PROSITE; PS01122; CASPASE_CYS; 1.
DR   PROSITE; PS01121; CASPASE_HIS; 1.
DR   PROSITE; PS50207; CASPASE_P10; 1.
DR   PROSITE; PS50208; CASPASE_P20; 1.
DR   PROSITE; PS50824; DAPIN; 1.
SQ   SEQUENCE   404 AA;  46077 MW;  629B8CFPB4A87D3B CRC64;

Query Match             10.5%;  Score 104;  DB 13;  Length 404;
Best Local Similarity   30.8%;  Pred. No. 0.38;
Matches   36;  Conservative   17;  Mismatches   50;  Indels   14;  Gaps    4;

Qy   10 DAIENLTAEELKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLETYGAELT 69
         |:|::| |:|||:|:  || || |  :  ||||| |     || || |  |||| |  |
Db   10 DVIEDLVEAELKQF-TRQLWIGVKPGVPEIPRGKLIENKDRQDVVDSMVQQYSEDAGT-IT 67

Qy   70 ANVLRDMGLQEMAGQL----------QAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQ 117
Db   68 VQTLRKIQNERAKRLESNLLKVQSQGQENKQNSBEP---QIPQIISQPIQQIISQ 121


RESULT 8
ID   Q8VPY3      PRELIMINARY;      PRT;   127 AA.
AC   Q8VPY3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SILE (Fragment).
OS   Serratia marcescens.
OG   Plasmid pR478.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21602765; PubMed=11739772;
RA   Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT   "Diversity of silver resistance genes in IncH incompatibility group
RT   plasmids.";
RL   Microbiology 147:3393-3402(2001).
DR   EMBL; AY009377; AAL68934.1; -.
KW   Plasmid.
FT   NON_TER      127        127
SQ   SEQUENCE   127 AA;  13438 MW;  67B3822C25BFC9A5 CRC64;

Query Match             9.2%;  Score 91.5;  DB 2;  Length 127;
Best Local Similarity   31.6%;  Pred. No. 1.1;
Matches   31;  Conservative   11;  Mismatches   47;  Indels    9;  Gaps    4;

Qy   64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIDQHRAAL 122
Db   12 FGLISSAWATETVNIHERVNNAQAPAHQMQSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

Qy   123 IARVTNVEWLLDALYGYKVLTDEQYQ--AVRAEPTNPSK 158
Db   67 IAHETMTNGSADA-HQKMVESHQRMIGSQTVSPTGPSK 103


RESULT 9
ID   Q8VPW8      PRELIMINARY;      PRT;   127 AA.
AC   Q8VPW8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SILE (Fragment).
OS   Salmonella oranienberg.
OG   Plasmid MIP225.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=28147;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21602765; PubMed=11739772;
RA   Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT   "Diversity of silver resistance genes in IncH incompatibility group
RT   plasmids.";
RL   Microbiology 147:3393-3402(2001).
DR   EMBL; AY009392; AAL68944.1; -.
KW   Plasmid.
FT   NON_TER      127        127
SQ   SEQUENCE   127 AA;  13456 MW;  67B3822C2313A365 CRC64;

Query Match             9.2%;  Score 91.5;  DB 2;  Length 127;
Best Local Similarity   31.6%;  Pred. No. 1.1;
Matches   31;  Conservative   11;  Mismatches   47;  Indels    9;  Gaps    4;

Qy   64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIDQHRAAL 122
Db   12 FGLISSAWATETVNIHERVNNAQAPAHQMQSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

Qy   123 IARVTNVEWLLDALYGYKVLTDEQYQ--AVRAEPTNPSK 158
Db   67 IAHETMTNGSADA-HQKMVESHQRMIGSQTVSPTGPSK 103


RESULT 10
ID   Q8VPX8      PRELIMINARY;      PRT;   127 AA.
AC   Q8VPX8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SILE (Fragment).
OS   Salmonella enterica subsp. enterica serovar Ohio.
OG   Plasmid pMIP233.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=117541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21602765; PubMed=11739772;
RA   Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT   "Diversity of silver resistance genes in IncH incompatibility group
RT   plasmids.";
RL   Microbiology 147:3393-3402(2001).
DR   EMBL; AY009382; AAL68941.1; -.
KW   Plasmid.
FT   NON_TER      127        127
SQ   SEQUENCE   127 AA;  13456 MW;  67B3822C2313A365 CRC64;

Query Match             9.2%;  Score 91.5;  DB 2;  Length 127;
Best Local Similarity   31.6%;  Pred. No. 1.1;
Matches   31;  Conservative   11;  Mismatches   47;  Indels    9;  Gaps    4;

Qy   64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIDQHRAAL 122
Db   12 FGLISSAWATETVNIHERVNNAQAPAHQMQSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

Qy   123 IARVTNVEWLLDALYGYKVLTDEQYQ--AVRAEPTNPSK 158
Db   67 IAHETMTNGSADA-HQKMVESHQRMIGSQTVSPTGPSK 103


RESULT 11
ID   Q8VPY8      PRELIMINARY;      PRT;   127 AA.
AC   Q8VPY8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

```
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SilE (Fragment).
GN   SilE.
OS   Serratia marcescens.
OG   Plasmid pR476b.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21602765; PubMed=11739772;
RA   Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT   "Diversity of silver resistance genes in IncH incompatibility group
RT   plasmids.";
RL   Microbiology 147:3393-3402 (2001).
KW   Plasmid.
DR   EMBL; AY009372; AAL68931.1; -.
FT   NON_TER      127      127
SQ   SEQUENCE   127 AA;  13456 MW;  67B3822C2313A365 CRC64;

Query Match             9.2%;  Score 91.5;  DB 2;  Length 127;
Best Local Similarity   31.6%;
Matches   31;  Conservative   11;  Mismatches   47;  Indels    9;  Gaps    4;

Qy   64  YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGGAPAGIQAPPQSAAKPGLHFIDQHRAAL 122
             :::::: :::::::::::::::::: ::  :: :::::::::::::::::::::::::: :
Db   11  FGLISSAWATETVNIHERVNNAQAPAHQMQSAAAPVGIQ----GTAPRMAGMQHEQAI  66

Qy  123  IARVTNVEWLLDALYGKVLTDEQYQ--AVRAEPTNPSK 158
             :::::  : :: :: : ::        :: :: :::
Db   67  IAHETMTNGSADA-HQKMVESHQRMMGSQTVSPTGPSK 103
```

```
RESULT 12
ID   Q9I9L7                 PRELIMINARY;      PRT;    383 AA.
AC   Q9I9L7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Caspase.
GN   CASPA OR CASPY.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20373792; PubMed=10917738;
RA   Inohara N., Nunez G.;
RT   "Genes with homology to mammalian apoptosis regulators identified in
RT   zebrafish.";
RL   Cell Death Differ. 7:509-510(2000).
DR   EMBL; AF233434; AAF66964.1; -.
HSSP; P29466; 1ICE.
ZFIN; ZDB-GENE-000616-3; caspa.
InterPro; IPR002398; ICE.
InterPro; IPR002138; ICE_p10.
InterPro; IPR001309; ICE_p20.
InterPro; IPR004020; PAAD_DAPIN_dom.
Pfam; PF00656; ICE_p20; 1.
Pfam; PF00656; ICE_p10; 1.
Pfam; PF02758; PAAD_DAPIN; 1.
PRINTS; PR00376; IL1BCENZYME.
SMART; SM00115; CASC; 1.
PROSITE; PS00376; CASPASE_CYS; 1.
PROSITE; PS01121; CASPASE_HIS; 1.
PROSITE; PS01122; CASPASE_HIS; 1.
PROSITE; PS50207; CASPASE_P10; 1.
PROSITE; PS50208; CASPASE_P20; 1.
PROSITE; PS50824; DAPIN; 1.
SEQUENCE   383 AA;  43966 MW;  2189087130977AC3 CRC64;

Query Match             9.2%;  Score 91.5;  DB 13;  Length 383;
Best Local Similarity   26.2%;
Matches   28;  Conservative   24;  Mismatches   44;  Indels   11;  Gaps    4;

Qy    5  RDAILDALENILTAEELKKFKLKLLSVPLREGYGRIPRGALLSM-DALDLTDKLVSFYLET  63
             :::::: :::: : : ::   :  :::::: :  :::::::::::  :::::::::::::::
Db    6  KDHLQDALSNIGADNLRRFQSRLGD---RKQEPRVRKSTIEKLKDEIDLVDLLVNTFTSD  62

Qy   64  YGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKP 110
             :: :::: ::::::::::::::::::::::::::::::::::::::::
Db   63  -AVSVTVDILRGIKCNAVAEELLENTGQG------GVSQPEPPVPEP 102
```

```
RESULT 13
ID   P74033                 PRELIMINARY;      PRT;    439 AA.
AC   P74033;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein slr0806.
GN   SLR0806.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90911; BAA18106.1; -.
DR   InterPro; IPR006094; Oxid_FAD_bind.
DR   Pfam; PF01565; FAD_binding_4; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   439 AA;  47989 MW;  6C2E6F56365FA653 CRC64;

Query Match             8.7%;  Score 86.5;  DB 16;  Length 439;
Best Local Similarity   24.0%;
Matches   42;  Conservative   22;  Mismatches   60;  Indels   51;  Gaps    8;

Qy   31  PLREGYGR---IPRGALLSMDALDLTDKLVSFYLETYGAELTANVLRDMGLQEMAGQLQA  87
             :::::: :    :: :::  : :: :  : :: ::   :::::::::::::::::::::::
Db  194  PVIHGYGTNGIITEITLPLTPALPWRRAIVSP-----------------TNLSSAIAPAQN 237

Qy   88  ATHQ-GSGAAPAGIQAPP----QSAAKPGLHFIDQHRAALIARVTNVEW----LL 133
             :::: :::::::::::::    ::::::::::::::::::::::::::::    ::
Db  238  LAHQDGIVSKEISIQADPIPQYFSSLKSYYQPGAHWV------MVIVSELDWLAFYQLA 290

Qy  134  DALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWWTCKDLLLQALRESQSYL 188
             :::::::::: :: :: : :: :  :: :::  :   ::: :: :::::: :::
Db  291  KASKGEIIFEQ---------DPQSPGKKINLEF---NWNHTLIARAVDPSLTYL 334
```

```
RESULT 14
ID   Q8NQN6                 PRELIMINARY;      PRT;    828 AA.
AC   Q8NQN6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Phenylalanyl-tRNA synthetase beta subunit (EC 6.1.1.20).
GN   CGL1390.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005278; BAB98783.1; -.
DR   InterPro; IPR005146; B3_4.
DR   InterPro; IPR005147; B5.
DR   InterPro; IPR005121; Fdx-AntiCB.
DR   InterPro; IPR004532; PheT_bact.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF03483; B3_4; 1.
DR   Pfam; PF03484; B5; 1.
DR   Pfam; PF03147; FDX-ACB; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   TIGRFAMs; TIGR00472; pheT_bact; 2.
DR   Pfam; PF01588; tRNA_bind; 1.
KW   Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ   SEQUENCE   828 AA;  88611 MW;  E3F1B39CE5FF29FE CRC64;
```

```
Query Match             8.7%;  Score 86.5; DB 16; Length 828;
Best Local Similarity  24.5%;  Pred. No. 41;
Matches  40;  Conservative  29; Mismatches  73; Indels  21; Gaps  7;

Qy   29  SVPLREGYGRIPRGALLSMDALDLTDKLVSFYL---ETYGAELTANVLRDMGLQEMAGQL   85
             :||||||||||||||||||||||||||||||||   ::|||| ::: ||  |||| ||
Db  107  AISARETYGRMSAGMICSASELGLADKQNSGIITLDPSYG-EPGEDARQALGLEDTVFDV  165

Qy   86  QAATHQSGGAAPAGIQAPPQSAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLT---  142
            |:|  ||  ||||| : |||| |||:|: |  ::::|||||:|||: |||||:|||
Db  166  NVTPDRGYALSARGLTRELASAF--SLTFTD---PAIEPAVAGIEVKVPAVEGSLINVEL  220

Qy  143  DEQYQAVRAEPTNPSKMRKLFSFTPAWN----WTCKDLLLQALR  182
            :|:|::|:|::|  ||:|:::|:||:|||    ||:|:|:
Db  221  REETKAIR------FGLRKVSGIDPAAESPFWMQREIMLSGQR  257
```

```
DR   Pfam; PF04675; DNA_ligase_A_N; 1.
DR   TIGRFAMs; TIGR00574; dnlj; 1.
DR   PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR   PROSITE; PS50160; DNA_LIGASE_A3; 1.
DR   PROSITE; PS00435; DNA_LIGASE_A1; 1.
DR   PROSITE; PS00435; PEROXIDASE_1; 1.
KW   Ligase; Complete proteome.
SQ   SEQUENCE   580 AA;  65351 MW;  31C511F878C79E73 CRC64;
```

```
Query Match             8.7%;  Score 86; DB 17; Length 580;
Best Local Similarity  23.1%;  Pred. No. 28;
Matches  45;  Conservative  30; Mismatches  60; Indels  60; Gaps  9;

Qy   10  DALENLTAEBLKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLIETYGAELT   69
             |:  ||| |||:|| |: ||||||||||||||||||||||||||||||||| |||||||
Db  136  DILQKATPEE-GKYIIRIVLGRLRLGFGD-----------------------QFLLEAFSIAFT  175

Qy   70  A-----------NVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPQSAAKPGLHFIDQ  117
            :|   |::|| | |||   |  ::  | :
Db  176  GDKXHAGKIKESYSVCTDIG--ELA---QTLAEHGAG------APGYPSIKPGRPV--  220

Qy  118  HRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDL-  176
            :: ||||||:|||:|||||||||:|:|||::||::::|::|:::|||:|||::|||||
Db  221  -KSMLAQRVESFEELEERIKGKKAAAEEKYDGERVQ-------IHKAGDEIKAFSRRLEDIT  273

Qy  177  ----LLQALREGQS  186
            ||:|||:|||:
Db  274  AQYPDIIEAVRESIS  288
```

```
Search completed: January 29, 2004, 13:48:13
Job time : 12.8067 secs
```

```
RESULT 15
ID   Q8TSR7      PRELIMINARY;      PRT;      580 AA.
AC   Q8TSR7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   DNA ligase (ATP).
GN   LIG OR MA0728.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21927760; PubMed=11932238;
RA   Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
DR   EMBL; AE010734; AAM04168.1; -.
DR   InterPro; IPR000977; Peroxidase.
DR   InterPro; IPR002016; Peroxidase.
DR   Pfam; PF01068; DNA_ligase_A; 1.
DR   Pfam; PF04679; DNA_ligase_A_C; 1.
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: January 29, 2004; 13:11:31 ; Search time 3.84236 Seconds
2147.276 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTAEEL..........LILQALRESQSYLVEDLRS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/pCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	990	100.0	195	4	US-09-340-620A-49	Sequence 49, Appl
2	684	69.1	193	4	US-09-340-620A-61	Sequence 61, Appl
3	378	38.2	71	4	US-09-340-620A-58	Sequence 58, Appl
4	264	26.7	70	4	US-09-340-620A-57	Sequence 57, Appl
5	264	26.7	70	4	US-09-340-620A-66	Sequence 66, Appl
6	151.3	15.3	109	4	US-09-340-620A-71	Sequence 71, Appl
7	94.5	9.5	1233	4	US-09-328-352-7874	Sequence 7874, Ap
8	86.5	8.7	372	1	US-08-813-584A-3	Sequence 3, Appli
9	86.5	8.7	372	1	US-08-330-515-3	Sequence 3, Appli
10	84.5	8.5	284	4	US-09-069-023-5	Sequence 5, Appli
11	84.5	8.5	478	4	US-09-069-023-4	Sequence 4, Appli
12	84.5	8.5	530	4	US-09-069-023-3	Sequence 3, Appli
13	84.5	8.5	531	4	US-09-069-023-1	Sequence 1, Appli
14	84.5	8.5	540	3	US-09-019-942-1	Sequence 1, Appli
15	84.5	8.5	540	4	US-09-099-041A-2	Sequence 2, Appli
16	84.5	8.5	540	4	US-09-069-023-27	Sequence 27, Appl
17	84.5	8.5	540	4	US-09-245-281-2	Sequence 2, Appli
18	84.5	8.5	540	4	US-09-470-271-1	Sequence 1, Appli
19	84.5	8.5	540	4	US-09-207-359B-2	Sequence 2, Appli
20	84.5	8.5	540	4	US-09-340-620A-2	Sequence 2, Appli
21	84.5	8.5	540	4	US-09-345-473E-28	Sequence 28, Appl
22	78.5	7.9	424	4	US-09-107-532A-6238	Sequence 6238, Ap
23	77.5	7.8	2436	3	US-08-044-818-75	Sequence 75, Appl
24	76.5	7.7	164	4	US-09-245-281-41	Sequence 41, Appl
25	76.5	7.7	164	4	US-09-207-359B-41	Sequence 41, Appl
26	76.5	7.7	164	4	US-09-245-281-39	Sequence 39, Appl
27	76.5	7.7	249	4	US-09-340-620A-39	Sequence 39, Appl
28	76.5	7.7	249	4	US-09-207-359B-39	Sequence 39, Appl
29	76.5	7.7	249	4	US-09-340-620A-39	Sequence 39, Appl
30	76.5	7.7	409	4	US-09-207-359B-46	Sequence 46, Appl
31	76.5	7.7	953	4	US-09-099-041A-8	Sequence 8, Appli
32	76.5	7.7	953	4	US-09-245-281-8	Sequence 8, Appli
33	76.5	7.7	953	4	US-09-340-620A-8	Sequence 8, Appli
34	76	7.7	435	4	US-09-252-991A-31215	Sequence 31215, A
35	76	7.7	1175	4	US-09-252-991A-25044	Sequence 25044, A
36	75.5	7.6	167	4	US-09-069-023-6	Sequence 6, Appli
37	75.5	7.6	3011	3	US-08-811-566-20	Sequence 20, Appl
38	75.5	7.6	3011	3	US-08-014-416-1	Sequence 1, Appli
39	75.5	7.6	3011	3	US-09-034-756-20	Sequence 20, Appl
40	75.5	7.6	3011	4	US-09-034-756-20	Sequence 20, Appl
41	75.5	7.6	3012	3	US-08-811-566-2	Sequence 2, Appli
42	75.5	7.6	3012	4	US-09-034-756-2	Sequence 2, Appli
43	75.5	7.6	3012	4	US-09-252-991A-32273	Sequence 32273, A
44	75	7.6	323	4	US-09-328-352-5345	Sequence 5345, Ap
45	75	7.6	372	4	US-09-328-352-5345	Sequence 5345, Ap

ALIGNMENTS

RESULT 1
US-09-340-620A-49
; Sequence 49, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREO
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-49

Query Match 100.0%; Score 990; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.8e-112; Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLRREGYGRIPRGALLSMDALDLTDKLVSFY 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLRREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy 121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy 181 LRESQSYLVEDLRS 195
         ||||||||||||||
Db 181 LRESQSYLVEDLRS 195

RESULT 2
US-09-340-620A-61
; Sequence 61, Application US/09340620A

; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-61

Query Match   69.1%;
Best Local Similarity   71.8%;
Matches 140; Conservative 19; Mismatches 34; Indels 0; Gaps 2;

Score 684; DB 4; Length 193;
Pred. No. 3.5e-75;

Qy   1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        |||||||||||||:|||||:||:|||||||||:||||||||||:|||||||||||||||||:|
Db   1 MGRARDAILDALENLSGDELKKFKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |:|| |:||:|| | |||||| : :|||||||||: ||||||||||||||
Db   61 LESYGLELTMTVLRDMGLQELAEQLO-TTKEESGAVAAAASVPAQSTARTG-HFVDQHRQ 118

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWN 180
        |||||| |::|||||| ||||:|||||||||||| |||||||||||||||
Db   119 ALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDXMRKLFSFVPSWNLTCKDSLLQA 178

Qy   181 LRESQSYLVEDLERS 195
        |:||| ||::||:|
Db   179 LKEIHPYLVMDLEQS 193

; RESULT 3
US-09-340-620A-58
; Sequence 58, Application US/0934o620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-58

Query Match   38.2%;
Best Local Similarity   100.0%;
Matches 71; Conservative 0;

Score 378; DB 4; Length 71;
Pred. No. 1.5e-38; Mismatches 0; Indels 0; Gaps 0;

Qy   111 GLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWN 170
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWN 60

Qy   171 WTCKDLLLQAL 181
        |||||||||||
Db   61 WTCKDLLLQAL 71

; RESULT 4
US-09-340-620A-57
; Sequence 57, Application US/0934o620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-57

Query Match   26.7%;
Best Local Similarity   73.9%;
Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Score 264; DB 4; Length 70;
Pred. No. 1.2e-24;

Qy   113 HFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWT 172
        ||:|||||||| ||||:|||||| ||||:|||||||||||| |||||||||||||||
Db   2 HFVDQHRQALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDXMRKLFSFVPSWNLT 61

Qy   173 CKDLLLQAL 181
        |:|| |||
Db   62 CKDSLLQAL 70

; RESULT 5
US-09-340-620A-66
; Sequence 66, Application US/0934o620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus

US-09-340-620A-66

```
Query Match             26.7%;   Score 264;  DB 4;  Length 70;
Best Local Similarity   73.9%;   Pred. No. 1.2e-24;
Matches  51;  Conservative   7;  Mismatches  11;  Indels    0;  Gaps    0;

Qy    113 HFIDQHRAALIARVTNVEWLLDALIYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWNT 172
              | :|:||:|||:|| || :|||||||| :|||::|::|||||||||||||||| :|:|
Db      2 HFVDQHRQLLIARVTEVDGVLDALLHGSVLTEGQYQAVRARTTSGQDKMRKLFSFVPSWNLT 61

Qy    173 CKDLLLQAL 181
              |||||||||
Db     62 CKDSLLQAL 70
```

RESULT 6
US-09-340-620A-71
; Sequence 71, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR APPLICATION FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR APPLICATION FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR APPLICATION FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR APPLICATION FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; NAME/KEY: VARIANT
; LOCATION: (1)...(109)
; OTHER INFORMATION: Xaa = Any Amino Acid

```
Query Match             15.3%;   Score 151.5;  DB 4;  Length 109;
Best Local Similarity   46.7%;   Pred. No. 1.2e-10;
Matches  42;  Conservative  13;  Mismatches  28;  Indels    7;  Gaps    4;

Qy    111 GLHFIDQHRAALIARVTN--VEWLLDALYGK-VLTDEQYQAVRAEPTNPSKMRKLFSFTP 167
              |||||||: :|||||| | :||||:|||| |:|||||||||||||||||||||:::||
Db      7 GSRIIDQHRXALLARVTEDPXDSLLDALLSRDLISEEDYBAVRAETTXLSKVRCLLILVQ 66

Qy    168 A-WWWTCKDL---LLQALRESQSYLVEDLE 193
              | |||     ::||||:||| :||||:|||
Db     67 SKGEETCXKFLKCLLQALKDSAAIGLDPE 96
```

RESULT 7
US-09-328-352-7874
; Sequence 7874, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7874
; LENGTH: 1233

; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7874

```
Query Match              9.5%;   Score 94.5;  DB 4;  Length 1233;
Best Local Similarity   26.5%;   Pred. No. 0.047;
Matches  53;  Conservative  26;  Mismatches  76;  Indels   45;  Gaps   10;

Qy    113 HFIDQHRAALIARVTNVE-WLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNW 171
              ||:|||||||||||||:  |||:||||||||||||||:|:|||:|||||||||||:||: |
Db    171 LIHDVLREWTQSQPQTVIDALYLAGELKSVDSFVKLVEDSLNFSSAAHFKLP---EKPTI 226

Qy    172 TCKDLLLQALRESQSYLVED 191
              |:|:|||||:||||||||||
Db    227 QF--EQLAQLQLATEIDISLLEPYI--LLDGEHYKHVNGTIFRNGAFNKLFS------ 275
```

RESULT 8
US-07-813-584A-3
; Sequence 3, Application US/07813584A
; Patent No. 5352588
; GENERAL INFORMATION:
; APPLICANT: Fischetti, Vincent A.
; APPLICANT: Bessen, Debra E
; TITLE OF INVENTION: No. 535258881 Immunoglobulin A Binding Protein
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kittie Murray
; STREET: 98 Cutter Mill Road
; CITY: Great Neck
; STATE: NY
; COUNTRY: USA
; ZIP: 11021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,584A
; FILING DATE: 19911224
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Kittie
; REGISTRATION NUMBER: 30,246
; REFERENCE/DOCKET NUMBER: RU-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-482-1990
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-813-584A-3

```
Query Match              8.7%;   Score 86.5;  DB 1;  Length 372;
Best Local Similarity   26.8%;   Pred. No. 0.066;
Matches  53;  Conservative  30;  Mismatches  78;  Indels   37;  Gaps    9;

Qy    113 HFIDQHRAALIARVTNVE-WLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNW 171
              ||:|||||||||||||:  |||:|:||||||||||::|:|||:||:||:|:|::||::|
Db     68 LTANVLRD-------MGLQEMAGQLQAA----THQSGAAPAGIQAPPQSAAKPGL 112

Qy    172 TCKDLLLQALRESQSYLVED 191
              |:|||||||:|||||||||
Db    125 VIDQLDELFVGTLDSFSQKLIREFAFES-GKIER-AQITDDA----------KTYSRQ 170
```

Left column

```
Db  134 EASRKGLSRDLEASRTAKELEAHQKLEAENKLTEGNQVSEASRKGLSNDLEASRAA 192
Qy  123 IARVTNVEWLLDALYGKVLTDEQ------YQAVRAEPTNPSKMRKLFSFTPAWWTCKDLL 177
Db  193 ------KKELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLS------RD-- 234
Qy  178 LQALRESQSYLVEDLERS 195
Db  235 LEASREANKKVTSELTQA 252
```

RESULT 9
US-08-330-515-3
Sequence 3, Application US/08330515
Patent No. 5556944

; GENERAL INFORMATION:
;   APPLICANT: Fischetti, Vincent A.
;   APPLICANT: Bessen, Debra E.
;   TITLE OF INVENTION: No. 5556944el Immunoglobulin A Binding Protein
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Saliwanchik & Saliwanchik
;   STREET: 2421 N.W. 41st Street, Suite A-1
;   CITY: Gainesville
;   STATE: FL
;   COUNTRY: USA
;   ZIP: 32606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/330,515
;   FILING DATE:
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/813,584
;   FILING DATE: 24-DEC-1991
;   CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Saliwanchik, David R.
;   REGISTRATION NUMBER: 31,794
;   REFERENCE/DOCKET NUMBER: RU-100.D1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 904-375-8100
;   TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 372 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-330-515-3

Query Match   8.7%;   Score 86.5;   DB 1;   Length 372;
Best Local Similarity   26.8%;   Pred. No. 0.068;
Matches   53;   Conservative   30;   Mismatches   78;   Indels   37;   Gaps   9;

```
Qy   12 LENLTAEELKKFKLKLLSVPLREGYGRIRPGALLSMDAL------DLTDKLVSFYLETYGA  66
Db   67 ELTANVL-RDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQ--SAAKPGL-HFIDQHRAAL 122
Qy   78 LEKINABEENKLBAINKBLNENYYKLQDG----IDALEKEKEDLKTTLAKTTKENEIS   133
Db  134 EASRKGLSRDLEASRTAKELEAHQKLEAENKLTEGNQVSEASRKGLSNDLEASRAA-   192
Qy  123 IARVTNVEWLLDALYGKVLTDEQ------YQAVRAEPTNPSKMRKLFSFTPAWWTCKDLL 177
Db  193 ------KKELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLS------RD--   234
Qy  178 LQALRESQSYLVEDLERS 195
```

Right column

```
Db  235 LEASREANKKVTSELTQA 252
```

RESULT 10
US-09-069-023-5
Sequence 5, Application US/09069023A
Patent No. 6348573

; GENERAL INFORMATION:
;   APPLICANT: Nunez, Gabriel
;   APPLICANT: Inohara, Naohiro
;   APPLICANT: Koseki, Takeyoshi
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
;   TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
;   FILE REFERENCE: UM-03333
;   CURRENT APPLICATION NUMBER: US/09/069,023A
;   CURRENT FILING DATE: 1998-04-27
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-069-023-5

Query Match   8.5%;   Score 84.5;   DB 4;   Length 284;
Best Local Similarity   20.1%;   Pred. No. 0.079;
Matches   50;   Conservative   33;   Mismatches   71;   Indels   95;   Gaps   9;

```
Qy    6 DAILDALENLTAEE-----LKKFKLKLLSVPLREGYGRIRPGALLSMDALDLTDKLVSFYL  61
Db   35 EPVLRTFBEITFLEAVIQLKKTKLQSVS------------------SAIHLCDK------  70
Qy   62 ETYGAELTANVLRDMGLQ-EMAGQLQAATHQGSGAAPAGIQAP------------------ 103
Db   71 --KKMELSLNIPVNHGPQEBSCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMK 128
Qy  104 ----------------------PQSAA-------KFGL--HFIDQHR 119
Db  129 LHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKR 188
Qy  120 AALIARVTN--VEWLLDALYGK-VLTDEQYQAVRAEPTNPSKMRKLFSFTPAWWTCKDL  176
Db  189 EDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSVKRQLLDTTDIQGEEFAKV 248
Qy  177 LIQALRESQ 185
Db  249 IVQKLKDNK 257
```

RESULT 11
US-09-069-023-4
Sequence 4, Application US/09069023A
Patent No. 6348573

; GENERAL INFORMATION:
;   APPLICANT: Nunez, Gabriel
;   APPLICANT: Inohara, Naohiro
;   APPLICANT: Koseki, Takeyoshi
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
;   TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
;   FILE REFERENCE: UM-03333
;   CURRENT APPLICATION NUMBER: US/09/069,023A
;   CURRENT FILING DATE: 1998-04-27
;   NUMBER OF SEQ ID NOS: 38
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 478
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-069-023-4

Query Match   8.5%;   Score 84.5;   DB 4;   Length 478;
Best Local Similarity   20.1%;   Pred. No. 0.18;

RESULT 12
US-09-069-023-3
; Sequence 3, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Nachiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 3
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-3

Query Match   8.5%; Score 84.5; DB 4; Length 530;
Best Local Similarity   20.1%; Pred. No. 0.21;
Matches   50; Conservative   33; Mismatches   71; Indels   95; Gaps   9;

Qy   6 DAILDALENLITAEB----LKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYL 61
            |:|| ::|: :|:::|| ||:|:|:
Db   229 EPVLRTFEEITFLEAVIQLKKTKLQSVS-----------------------SAIHLCDK----- 264

Qy   62 ETYGAELTANVLRDMGLQ--EMAGQLQAATHQGSGAAPAGIQAP-------------------- 103
            |||: ||| |: :: :|
Db   265 --KKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMK 322

Qy   104 --POSAA------KPGL--HFIDQHR 119
                                                          |: ||: | :
Db   323 LHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKR 382

Qy   120 AALIARVTN--VEWLLDALYGK-VLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDL 176
            :||| ::: :|:| || :::| |: | :
Db   383 EDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKV 442

Qy   177 LIQALRESQ 185
            ::| |:::
Db   443 IVQKLKDNK 451

RESULT 13
US-09-023-1
; Sequence 1, Application US/09069023A
; Patent No. 6348573

RESULT 14
US-09-019-942-1
; Sequence 1, Application US/09019942
; Patent No. 6033855
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
; TITLE OF INVENTION: DOMAIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,942
; FILING DATE: 06-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/068001

Query Match   8.5%; Score 84.5; DB 4; Length 531;
Best Local Similarity   20.1%; Pred. No. 0.21;
Matches   50; Conservative   33; Mismatches   71; Indels   95; Gaps   9;

Qy   6 DAILDALENLITAEB----LKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYL 61
            |:|| ::|: :|:::|| ||:|:|:
Db   282 EPVLRTFEEITFLEAVIQLKKTKLQSVS-----------------------SAIHLCDK----- 317

Qy   62 ETYGAELTANVLRDMGLQ--EMAGQLQAATHQGSGAAPAGIQAP-------------------- 103
            |||: ||| |: :: :|
Db   318 --KKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAPQDNDFLSRKAQDCYFMK 375

Qy   104 --POSAA------KPGL--HFIDQHR 119
                                                          |: ||: | :
Db   376 LHHCPGNHSWDSTISGSQRAAFCDHKTTPCSSAIINPLSTAGNSERLQPGIAQQWIQSKR 435

Qy   120 AALIARVTN--VEWLLDALYGK-VLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDL 176
            :||| ::: :|:| || :::| |: | :
Db   436 EDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKV 495

Qy   177 LIQALRESQ 185
            ::| |:::
Db   496 IVQKLKDNK 504

```
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 617/542-5070
;       TELEFAX: 617/542-8906
;       TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 540 amino acids
;           TYPE: amino acid
;           TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-09-019-942-1
```

```
Query Match     8.5%; Score 84.5; DB 3; Length 540;
Best Local Similarity 20.1%; Pred. No. 0.21;
Matches   50; Conservative  33; Mismatches  71; Indels  95; Gaps  9;

Qy    6 DAILDALENLITAEE------LKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYL 61
         |:|:|:|:||:|:||   : :| ||:||:|| ||:| :|:|:|::|:||::|:
Db  291 EPVLRTFEEITFLEAVIQLKKTKLQSVS------------------SAIHLCDK------ 326

Qy   62 ETYGAELTANVLRDMGLQ-EMAGLQAATHQGSGAAPAGIQAP---------------- 103
         ||:|:||  :|  |||::|||  :|   :::
Db  327 --KKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAFQDNDFLSRKAQDCYFMK 384

Qy  104 ------------------------PQSAA--------KPGL--HFIDQHR 119
                                  ||:||         :|:|  |:|:|:|
Db  385 LHHCPGNHSWDSTISGSQRAAFCDHKTIPCSSAIINPLSTAGNSERLQPGIAQQWIQSKR 444

Qy  120 AALIARVTN---VEWLLDALYGK-VLITDEQYQAVRAAEPTNPSKMRKLFSFTPAWWTCKDL 176
         ::|:| | |   ||||| |  :||  :|
Db  445 EDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKV 504

Qy  177 LLQALRESQ 185
         ::|:|:::
Db  505 IVQKLKDNK 513
```

```
RESULT 15
US-09-099-041A-2
Sequence 2, Application US/0909041A
Patent No. 6340576
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
    LENGTH: 540
    TYPE: PRT
    ORGANISM: Homo sapiens
US-09-099-041A-2
```

```
Query Match     8.5%; Score 84.5; DB 4; Length 540;
Best Local Similarity 20.1%; Pred. No. 0.21;
Matches   50; Conservative  33; Mismatches  71; Indels  95; Gaps  9;

Qy    6 DAILDALENLITAEE------LKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYL 61
         |:|:|:|:||:|:||   : :| ||:||:|| ||:| :|:|:|::|:||::|:
Db  291 EPVLRTFEEITFLEAVIQLKKTKLQSVS------------------SAIHLCDK------ 326

Qy   62 ETYGAELTANVLRDMGLQ-EMAGLQAATHQGSGAAPAGIQAP---------------- 103
         ||:|:||  :|  |||::|||  :|   :::
Db  327 --KKMELSLNIPVNHGPQEESCGSSQLHENSGSPETSRSLPAFQDNDFLSRKAQDCYFMK 384

Qy  104 ------------------------PQSAA--------KPGL--HFIDQHR 119
                                  ||:||         :|:|  |:|:|:|
Db  385 LHHCPGNHSWDSTISGSQRAAFCDHKTIPCSSAIINPLSTAGNSERLQPGIAQQWIQSKR 444
```

```
Qy  120 AALIARVTN---VEWLLDALYGK-VLITDEQYQAVRAAEPTNPSKMRKLFSFTPAWWTCKDL 176
         ::|:| | |   ||||| |  :||  :|
Db  445 EDIVNQMTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLLDTTDIQGEEFAKV 504

Qy  177 LLQALRESQ 185
         ::|:|:::
Db  505 IVQKLKDNK 513
```

Search completed: January 29, 2004, 13:48:52
Job time : 3.84236 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6

Run on: January 29, 2004, 13:46:01 ; Search time 33.9809 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTAEEL.........LILQALRESQSYLVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Published Applications AA:*
1: cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

789580

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID		Description
1	990	100.0	195	9	US-09-728-721-49	Sequence 49, Appli
2	990	100.0	195	10	US-09-996-617-8	Sequence 8, Appli
3	990	100.0	195	10	US-09-841-879B-5	Sequence 5, Appl
4	990	100.0	195	12	US-10-131-410-84	Sequence 84, Appl
5	990	100.0	195	12	US-10-240-145-55	Sequence 55, Appli
6	990	100.0	195	15	US-10-295-981-49	Sequence 49, Appli
7	990	100.0	205	15	US-10-106-698-5421	Sequence 5421, Appli
8	759	76.7	193	9	US-09-925-301-1120	Sequence 1120, Ap
9	684	69.1	193	9	US-09-728-721-61	Sequence 61, Appl
10	684	69.1	193	10	US-09-841-879B-2	Sequence 2, Appli
11	684	69.1	193	15	US-10-106-698-5422	Sequence 5422, Appl
12	564.5	57.0	158	15	US-10-295-981-61	Sequence 61, Appli
13	509	51.4	136	9	US-09-925-299-1421	Sequence 1421, Ap
14	509	51.4	136	11	US-09-925-299-1421	Sequence 1421, Ap
15	469	47.4	90	10	US-09-931-071-7	Sequence 7, Appli
16	445	44.9	85	10	US-09-841-879B-8	Sequence 8, Appli
17	435	43.9	89	14	US-10-127-516-16	Sequence 16, Appl
18	435	43.9	14	9	US-10-027-629-16	Sequence 9, Appli
19	378	38.2	71	9	US-09-728-721-58	Sequence 58, Appl
20	378	38.2	71	15	US-10-295-981-58	Sequence 58, Appl
21	378	38.1	77	12	US-10-132-516-8	Sequence 8, Appl
22	377	38.1	77	14	US-10-027-629-8	Sequence 8, Appli
23	370	37.4	77	12	US-10-132-967-8	Sequence 8, Appli
24	368	37.2	76	11	US-09-965-621-9	Sequence 9, Appli
25	368	37.2	76	12	US-10-407-866-9	Sequence 9, Appli
26	309	31.2	85	10	US-09-841-879B-15	Sequence 15, Appl
27	301	30.4	84	11	US-09-841-879B-7	Sequence 7, Appli
28	267	27.0	89	11	US-09-965-621-28	Sequence 28, Appl
29	267	27.0	89	12	US-10-407-866-28	Sequence 28, Appl
30	267	27.0	89	14	US-10-027-629-17	Sequence 17, Appli
31	267	27.0	89	14	US-10-132-516-17	Sequence 17, Appl
32	264	26.7	70	9	US-09-728-721-57	Sequence 57, Appl
33	264	26.7	70	15	US-10-295-981-57	Sequence 57, Appl
34	264	26.7	70	9	US-09-728-721-66	Sequence 66, Appl
35	264	26.7	70	15	US-10-295-981-66	Sequence 66, Appl
36	250.5	25.3	70	15	US-10-295-981-57	Sequence 57, Appl
37	250.5	25.3	70	9	US-09-931-071-2	Sequence 2, Appl1
38	250.5	25.3	70	14	US-09-996-617-2	Sequence 2, Appl1
39	250.5	25.3	70	10	US-09-388-221-4	Sequence 4, Appl1
40	250.5	25.3	70	12	US-10-183-770-15	Sequence 15, Appl
41	250.5	25.3	70	15	US-10-028-392-11	Sequence 11, Appl1
42	250	25.3	10	15	US-09-388-221-6	Sequence 6, Appl1
43	250	25.3	10	12	US-09-388-221-2	Sequence 2, Appl1
44	249.5	25.2	12	11	US-10-028-374-3	Sequence 3, Appl1
45	249.5	25.2	12	10	US-10-183-770-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-728-721-49
Sequence 49, Application US/09728721
Patent No. US20020061845A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/728,721
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: 09/340,620
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 49
LENGTH: 195
TYPE: PRT
ORGANISM: Homo sapiens
US-09-728-721-49

Query Match 100.0%; Score 990; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy  61 LETYGAELTANVLRDMGLQEWAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 LETYGAELTANVLRDMGLQEWAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
```

```
Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

RESULT 2
US-09-996-617-8
; Sequence 8, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
;   APPLICANT: Bertin, John
;   TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
;   TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
;   FILE REFERENCE: 07334-340001
;   CURRENT APPLICATION NUMBER: US/09/996,617
;   CURRENT FILING DATE: 2001-11-27
;   PRIOR APPLICATION NUMBER: 09/931,071
;   PRIOR FILING DATE: 1999-06-28
;   PRIOR APPLICATION NUMBER: 2001-08-15
;   PRIOR APPLICATION NUMBER: 09/428,252
;   PRIOR FILING DATE: 1999-10-27
;   PRIOR APPLICATION NUMBER: 09/340,620
;   PRIOR FILING DATE: 1999-06-28
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-617-8

Query Match   100.0%;   Score 990; DB 10;   Length 195;
Best Local Similarity   100.0%;   Pred. No. 3.1e-98;
Matches 195;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

RESULT 3
US-09-841-879B-5
; Sequence 5, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
;   APPLICANT: Bertin, John
;   APPLICANT: Bertin, John
;   TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;   FILE REFERENCE: 07334-330001
;   CURRENT APPLICATION NUMBER: US/09/841,879B
;   CURRENT FILING DATE: 2001-04-24
;   PRIOR APPLICATION NUMBER: US 09/728,721
;   PRIOR FILING DATE: 2000-12-01
;   PRIOR APPLICATION NUMBER: US 09/340,620
;   PRIOR FILING DATE: 1999-06-28
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-841-879B-5

Query Match   100.0%;   Score 990; DB 10;   Length 195;
Best Local Similarity   100.0%;   Pred. No. 3.1e-98;
Matches 195;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

RESULT 4
US-10-131-410-84
; Sequence 84, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
;   APPLICANT: SPECHT, THOMAS
;   APPLICANT: HINZMANN, BERND
;   APPLICANT: SCHMITT, ARMIN
;   APPLICANT: PILARSKY, CHRISTIAN
;   APPLICANT: DAHL, EDGAR
;   APPLICANT: ROSENTHAL, ANDRE
;   TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
;   TITLE OF INVENTION: TUMORS
;   FILE REFERENCE: SCH-1763
;   CURRENT APPLICATION NUMBER: US/10/131,410
;   CURRENT FILING DATE: 2002-04-25
;   PRIOR APPLICATION NUMBER: 09/646,673
;   PRIOR FILING DATE: 2000-09-20
;   PRIOR APPLICATION NUMBER: PCT/DE99/00908
;   PRIOR FILING DATE: 1999-03-19
;   NUMBER OF SEQ ID NOS: 202
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 84
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-410-84

Query Match   100.0%;   Score 990; DB 12;   Length 195;
Best Local Similarity   100.0%;   Pred. No. 3.1e-98;
Matches 195;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy  181 LRESQSYLVEDLERS 195
        |||||||||||||||
Db  181 LRESQSYLVEDLERS 195
```

RESULT 5
US-10-240-145-55
; Sequence 55, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 55
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145-55

Query Match   100.0%;   Score 990; DB 12; Length 195;
Best Local Similarity   100.0%;   Pred. No. 3.1e-98;
Matches 195; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy   181 LRESQSYLVEDLERS 195
              |||||||||||||||
Db   181 LRESQSYLVEDLERS 195

RESULT 6
US-10-295-981-49
; Sequence 49, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-981-49

Query Match   100.0%;   Score 990; DB 15; Length 195;
Best Local Similarity   100.0%;   Pred. No. 3.1e-98;
Matches 195; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180

Qy   181 LRESQSYLVEDLERS 195
              |||||||||||||||
Db   181 LRESQSYLVEDLERS 195

RESULT 7
US-10-106-698-5421
; Sequence 5421, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5421
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5421

Query Match   100.0%;   Score 990; DB 15; Length 205;
Best Local Similarity   100.0%;   Pred. No. 3.4e-98;
Matches 195; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 130

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 190

Qy   181 LRESQSYLVEDLERS 195
              |||||||||||||||
Db   191 LRESQSYLVEDLERS 205

RESULT 8
US-09-925-301-1120
; Sequence 1120, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1120

  Query Match   76.7%; Score 759; DB 9; Length 190;
  Best Local Similarity   95.6%; Pred. No. 2.1e-73;
  Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Qy   1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 130

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR 160
            ||||||||||||||||||||||||||||||||||||
Db   131 ALIARVTNVEWLLDALYGKVLTDEQYQAVR--PSPPTQAR 168

RESULT 9
US-09-728-721-61
; Sequence 61, Application US/09728721
; Patent No. US2002061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-728-721-61

  Query Match   69.1%; Score 684; DB 9; Length 193;
  Best Local Similarity   71.8%; Pred. No. 2.6e-65;
  Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy   1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
            |||||||||||||||||||:::|||||||:||||||||||||||||||||||||||:|
Db   1 MGRARDAILDALENLSGDELKKFKMKLITVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

RESULT 10
US-09-841-879B-2
; Sequence 2, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-841-879B-2

  Query Match   69.1%; Score 684; DB 10; Length 193;
  Best Local Similarity   71.8%; Pred. No. 2.6e-65;
  Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy   1 MGRARDAILDALENLTABELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
            |||||||||||||||||:::|||||||:||||||||||||||||||||||||||||:|
Db   1 MGRARDAILDALENLSGDELKKFKMKLITVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
            |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKEESGAVAAAASVPAQSTARTG-HFVDQHRQ 118

Qy   121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
            ||||||:|:|||||||||||||||:||:| |||||:||||||||||||||||||:||||
Db   119 ALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDKMRKLFSFVPSWNLTCKDSLLQA 178

Qy   181 LRESQSYLVEDLERS 195
            |:| ||||:||||:|
Db   179 LKEIHPYLVMDLEQS 193

RESULT 11
US-10-295-981-61
; Sequence 61, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREC
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041

```
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 61
  LENGTH: 193
  TYPE: PRT
  ORGANISM: Mus musculus
US-10-295-981-61

Query Match       69.1%;
Best Local Similarity   71.8%;   Score 684;  DB 15;  Length 193;
Matches  140;  Conservative  19;  Mismatches  34;  Indels  2;  Gaps  2;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
         ||||||||||||||:|:|||||||:||||:||:||||||||||||||:|||||||||:|:
Db    1 MGRARDAILDALENLSGDELKKFKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSYY 60

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRA 120
         ||:|||||:|:|||||||:|:|:-||-|::::::||:||:::::::|:|:||:|||||||:
Db   61 LESYGLELTMTVLRDMGLQELAEQL-Q-TYKEBSGAVAAAASVPAQSTARYG-HFVDQHRQ 118

Qy  121 ALIARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMRKLFSFTPAWNWTCKDLLLQA 180
         ||:||||:|:|:||||:|||||:|||||||:|:||:||||:|||||:|||||:||||||||
Db  119 ALIARVTEVDGVLDALHGSVLTEGQYQAVRAETTSQDRVRKLFSFVPSWNLTCKDSLLQA 178

Qy  181 LRESQSYLVEDLRRS 195
         |:|||||||||||:|
Db  179 LKEIHPYLVWDLEQS 193


RESULT 12
US-10-106-698-5422
Sequence 5422, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: Patentin Ver. 3.0
SEQ ID NO 5422
  LENGTH: 158
  TYPE: PRT
  ORGANISM: Homo sapiens
US-10-106-698-5422

Query Match       57.0%;
Best Local Similarity   75.8%;   Score 564.5;  DB 15;  Length 158;
Matches  119;  Conservative  3;  Pred. No. 1.4e-52;  Mismatches  2;  Indels  33;  Gaps  2;

Qy    4 ARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFYLET 63
         ||||||||||||||||||||||                                 ||||||
Db   13 ARDAILDALENLTAEELKF----------------------------------KLVSFYLET 41

Qy   64 YGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALI 123
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42 YGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAGIQAPPQSAAKPGLHFIDQHRAALI 101

Qy  124 ARVTNVEWLLDALYGKVLTDEQYQAVRAEPTNPSKMR 160
         |||||||||||||||||||||||||:|:::|
Db  102 ARVTNVEWLLDALYGKVLTDEQYQAVR--PSPPTQAR 136
```

```
RESULT 13
US-09-925-299-1421
Sequence 1421, Application US/09925299
Patent No. US2005005627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1421
  LENGTH: 136
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match       51.4%;
Best Local Similarity   83.7%;   Score 509;  DB 9;  Length 136;
Matches  108;  Conservative  2;  Pred. No. 1e-46;  Mismatches  11;  Indels  8;  Gaps  2;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAG----IQAPPQSAAKPGLHFID 116
         |||||||||||||||||||||||||||||||||||||||      |||||||||||||||
Db   71 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPLGSRPLLSRQPSQACT----LID 126

Qy  117 QHRAALIAR 125
         ||||:|
Db  127 QHRASLSRR 135


RESULT 14
US-09-925-299-1421
Sequence 1421, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1421
  LENGTH: 136
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match       51.4%;
Best Local Similarity   83.7%;   Score 509;  DB 11;  Length 136;
Matches  108;  Conservative  2;  Pred. No. 1e-46;  Mismatches  11;  Indels  8;  Gaps  2;

Qy    1 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 MGRARDAILDALENLTAEELKKFKLKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70

Qy   61 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPAG----IQAPPQSAAKPGLHFID 116
         |||||||||||||||||||||||||||||||||||||||      ||||||||||:||
Db   71 LETYGAELTANVLRDMGLQEMAGQLQAATHQGSGAAPLGSRPLLSRQPSQACT----LID 126
```

```
Qy    117 QHRAALIAR 125
          ||||:|
Db    127 QHRASLSRR 135
```

RESULT 15
US-09-931-071-7
; Sequence 7, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-7

Query Match   47.4%; Score 469; DB 10; Length 90;
Best Local Similarity   100.0%; Pred. No. 1.2e-42;
Matches   90; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   106 SAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRABPTNPSKMRKLFSF 165
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 SAAKPGLHFIDQHRAALIARVTNVEWLLDALYGKVLTDEQYQAVRABPTNPSKMRKLFSF 60

Qy   166 TPAWNWTCKDLLLQALRESQSYLVEDLERS 195
          |||||||||||||||||||||||||||||
Db   61 TPAWNWTCKDLLLQALRESQSYLVEDLERS 90
```

Search completed: January 29, 2004, 13:54:38
Job time : 34.9809 secs

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 19587.5 Seconds
(without alignments)
11370.118 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gcccaggggcctggagaggt.........aggaataagaagtcacctac 5444

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_vrt:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rod:*
37: em_htg_mam:*
38: em_htg_vrt:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| C | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 5444 | 100.0 | 5444 | 9 | AB023143 | AB023143 Homo sapi |
| | 2 | 5400.8 | 99.2 | 5404 | 6 | AF310105 | AF310105 Homo sapi |
| | 3 | 4882.4 | 89.7 | 5100 | 6 | AX459865 | AX459865 Sequence |
| | 4 | 4882.4 | 89.7 | 5100 | 9 | AX207661 | AX207661 Sequence |
| | 5 | 4765.6 | 87.5 | 5122 | 9 | AF229059 | AF229059 Homo sapi |
| | 6 | 4290 | 78.8 | 4290 | 9 | AF229059 | AF229059 Homo sapi |
| | 7 | 4281.8 | 78.7 | 5306 | 9 | BC051787 | BC051787 Sequence |
| | 8 | 4148 | 76.2 | 4422 | 9 | AX089761 | AX089761 Sequence |
| | 9 | 4148 | 76.2 | 4422 | 9 | AF229060 | AF229060 Homo sapi |
| | 10 | 4100 | 75.3 | 4200 | 9 | AF229061 | AF229061 Homo sapi |
| | 11 | 4078 | 74.9 | 4194 | 9 | AX089763 | AX089763 Sequence |
| | 12 | 3958 | 72.7 | 4332 | 9 | AF229062 | AF229062 Homo sapi |
| | 13 | 3952 | 72.6 | 4329 | 9 | AX089765 | AX089765 Sequence |
| | 14 | 3807.8 | 69.9 | 4123 | 9 | HSM800983 | HSM800983 Homo sapi |
| | 15 | 3400.4 | 62.5 | 4556 | 6 | AX089769 | AX089769 Sequence |
| | 16 | 3210.4 | 59.0 | 4466 | 6 | AX089771 | AX089771 Sequence |
| | 17 | 3094.2 | 56.8 | 3419 | 6 | AK026398 | AK026398 Homo sapi |
| c | 18 | 1710.8 | 31.4 | 152812 | 2 | AC090536 | AC090536 Homo sapi |
| c | 19 | 1710.8 | 31.4 | 176730 | 2 | AC055839 | AC055839 Homo sapi |
| c | 20 | 1566.4 | 28.8 | 2373 | 9 | AK026398 | AK026398 Homo sapi |
| c | 21 | 991.8 | 18.2 | 111862 | 2 | HSAC001237 | AC001237 Homo sapi |
| | 22 | 954 | 17.5 | 111862 | 2 | HSAC001237 | AC001237 Homo sapi |
| c | 23 | 742.8 | 13.6 | 153021 | 10 | AL663042 | AL663042 Mouse DNA |
| c | 24 | 738 | 13.6 | 174564 | 2 | AC121704 | AC121704 Rattus no |
| c | 25 | 738 | 13.6 | 250275 | 2 | AC095695 | AC095695 Rattus no |
| c | 26 | 728 | 13.4 | 221540 | 2 | AC026912 | AC026912 Mus muscu |
| c | 27 | 722 | 13.3 | 239196 | 2 | AC127967 | AC127967 Rattus no |
| | 28 | 701.6 | 12.9 | 96115 | 10 | AL662908 | AL662908 Mouse DNA |
| c | 29 | 695.2 | 12.8 | 180274 | 2 | AC027185 | AC027185 Mus muscu |
| | 30 | 603 | 11.1 | 254690 | 2 | AC095179 | AC095179 Rattus no |
| | 31 | 399.8 | 7.3 | 3827 | 9 | AY095146 | AY095146 Homo sapi |
| | 32 | 399.8 | 7.3 | 3827 | 9 | AY154467 | AY154467 Homo sapi |
| c | 33 | 399.8 | 7.3 | 69124 | 2 | AC069567 | AC069567 Mus muscu |
| | 34 | 398.2 | 7.3 | 3466 | 9 | AK095460 | AK095460 Homo sapi |
| | 35 | 398.2 | 7.3 | 3507 | 9 | BC028069 | BC028069 Homo sapi |
| | 36 | 398 | 7.3 | 3108 | 9 | AX417214 | AX417214 Sequence |
| | 37 | 398 | 7.3 | 3300 | 6 | AX459869 | AX459869 Sequence |
| | 38 | 398 | 7.3 | 3563 | 9 | AY116205 | AY116205 Homo sapi |
| | 39 | 398 | 7.3 | 3731 | 9 | AY116204 | AY116204 Homo sapi |
| | 40 | 397.4 | 7.3 | 3395 | 9 | AY116206 | AY116206 Homo sapi |
| | 41 | 395.8 | 7.3 | 3221 | 9 | AY116207 | AY116207 Homo sapi |
| c | 42 | 392.2 | 7.2 | 108555 | 2 | AC079221 | AC079221 Mus muscu |
| | 43 | 376 | 6.9 | 3828 | 9 | AF427617 | AF427617 Homo sapi |
| | 44 | 374.4 | 6.9 | 3237 | 9 | AY092033 | AY092033 Homo sapi |
| | 45 | 373 | 6.9 | 4170 | 9 | AF410477 | AF410477 Homo sapi |

ALIGNMENTS

RESULT 1
AB023143
LOCUS        AB023143      5444 bp    mRNA    linear   PRI 16-JUN-1999
DEFINITION   Homo sapiens mRNA for KIAA0926 protein, complete cds.
ACCESSION    AB023143
VERSION      AB023143.1  GI:4589483
KEYWORDS     .
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (sites)
AUTHORS      Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirosawa,M.,
             Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE        Prediction of the coding sequences of unidentified human genes.

Pred. No. is the number of results predicted by chance to have a

XIII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro

```
JOURNAL    DNA Res. 6 (1), 63-70 (1999)
MEDLINE    99246063
PUBMED     10231032
REFERENCE  2  (bases 1 to 5444)
AUTHORS    Ohara,O., Nagase,T. and Kikuno,R.
TITLE      Direct Submission
JOURNAL    Submitted (04-FEB-1999) Osamu Ohara, Kazusa DNA Research Institute,
           Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
           292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp; Tel:+81-438-52-3913,
           Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
source          1..5444
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="hh02962"
                /sex="male"
                /tissue_type="brain"
                /clone_lib="pBluescriptII SK plus"
                /dev_stage="adult"
gene            1..5444
                /gene="KIAA0926"
CDS             523..4812
                /gene="KIAA0926"
                /codon_start=1
                /product="KIAA0926 protein"
                /protein_id="BAA76770.1"
                /db_xref="GI:4589484"
                /translation="MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSGETPAQ
                PEKTSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQQEGAGHSPSPYSPSE
                PHLGSPSQPTSTAVLMPWIHIELPAGCTQGSERRVLRQLPDTSGRRREISASLLYQAL
                PSSPDHESPSQPSNAPTSTAVLGSWSGSPPQPSLAPREQEAPGTQWPLDETSGIYYTE
                IRERREKSEKGRPPWAAAVGTPPQAHTSLQPHHHPWMEPSVRESLCSTWPWKNEDFNQ
                KFTQLILLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFGPGDILQTQEPRIVILQG
                AAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTLLP
                EASFLITARTPALQNLIPSLEQARWEVLGFSESSSRKEYFRYFTDERQAIRAFRLVK
                SNKELWALCLVPWVSWLACTCLMQQWRKEKLTLTSKTTTILCLHYLAQACLQAQPLGP
                QLRDLCSLAAEGIWQKKTLFSPDDLRKGHLDGAIISTFLKMGILQEHPIPLSYSFIHL
                CFQEFFAAMSVYLEDEKGRGRKHSNCIIDLEKTLEAYGIHGLPGASTRFLLGLLSDEG
                ERBMENIFHCRLSOGRNLMQWWPSLQLLLQPHSLESLHCLVETRNKFLTQVMAHFEES
                MGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQ
                ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGNGLTEAEDCKD
                LAFGRLANQTLTELDLSFNVLTDAGAKHLCORLROPSSCKLQRLQVSCGLTSDCCQDL
                ASVLSASPSLKELDLQQNLDDVGVRLLCEGRHPACKLIRLGLDQTTLSDEMRQELR
                ALEQEKPOLLIFSRRKPSVMPTEGDLGTEGEMSNSTSSLKRORLGSRAASHVAQANLK
                LLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTEKLGTDDFWGPTGPVAT
                EVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCWDQFPLGEINPOHSMM
                VAGPLDIKAEPGAVEAVHLPHFVALQGGHVDTISLFQMAHFKEEGMLLEKPARVELHH
                IVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLIPSDCSIRKE
                LELCYRSPGEDQLPSEFVVGHLGSGIRRLQVRDKDETLVWEALVKPGDLMPATTLIPP
                ARIAVSPLDAPQLLHFVDQYRBEOLIARVTSVEVTVLDKLHGQVLSQBQYERVLAENTR
                PSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS"
```

```
BASE COUNT    1303 a    1525 c    1474 g    1142 t
ORIGIN
```

```
Qy   Query Match    100.0%;   Score 5444;   DB 9;   Length 5444;
     Best Local Similarity   100.0%;   Pred. No. 0;
     Matches 5444;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 GCCCCAGGGCCTGGAGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCCCGGGGCTCCACTCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCCCAGGGCCTGGAGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCCCGGGGCTCCACTCT 60

Qy     61 GGGGTTCTGAAAAGCCCATTCCCTGCTCTCGGGGCTCCTCCCACCCCACCTCTTCTCAGCCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGTTCTGAAAAGCCCATTCCCTGCTCTCGGGGCTCCTCCCACCCCACCTCTTCTCAGCCTT 120

Qy    121 GCAGCTCAAGGGTTGATCTCAGGAGGTCCAGGAGGTCCAGGAGGAGGGAAGAAATCTGAGGAACA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCAGCTCAAGGGTTGATCTCAGGAGGTCCAGGAGGTCCAGGAGGAGGGAAGAAATCTGAGGAACA 180

Qy    181 CAGAACAGTGAGCGGTTGCCCACACCCATCTCCCGTCACCACATCTCCCCTCACCCTCAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGAACAGTGAGCGGTTGCCCACACCCATCTCCCGTCACCACATCTCCCCTCACC.240-

Qy    241 CCTCCCTGCCTGGCCCGCCCTGGCCCTGACCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCTGCCTGGCCCGCCCTGGCCCTGACCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT 300

Qy    301 GTCTTGCAGGCCCCCTCCTGGGCTCCTCCCCTCCCCCGTGGCTTTTCCTTACCACTCCCCCTCTAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTCTTGCAGGCCCCCTCCTGGGCTCCTCCCCTCCCCCGTGGCTTTTCCTTACCACTCCCCCTCTAT 360

Qy    361 CGGGCGTCTATCTGTAGGTGGCCCTCGGGGATTTATAAAAACTGGGGTTCGGAATGCTGAATAAGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGGGCGTCTATCTGTAGGTGGCCCTCGGGGATTTATAAAAACTGGGGTTCGGAATGCTGAATAAGA 420

Qy    421 GACGGTAAGGAGGCCAAAGGCAAAAGGACAGACACTGTTCTCTCCCTGCTGATACCCTCACCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GACGGTAAGGAGGCCAAAGGCAAAAGGACAGACACTGTTCTCTCCCTGCTGATACCCTCACCAC 480

Qy    481 CTGGGAACATCCCCCAGACACCCTCTTAACTCCGGGACCAGAGACGAGGCTGGCGGGAGCCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGGGAACATCCCCCAGACACCCTCTTAACTCCGGGACCAGAGACGAGGCTGGCGGGAGCCTGG 540

Qy    541 GGCCGCCCTGGCCGTTACTTGGAGTTCCTGAAGGAGGAGGAGGCTGAAGGAGGTTCCAGCTT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGCCGCCCTGGCCGTTACTTGGAGTTCCTGAAGGAGGAGGAGGCTGAAGGAGGTTCCAGCTT 600

Qy    601 CTGCTCGCCAATAAAGGCGCACTCCAGGAGCTCTTCGGGGTGAGACACACCGCTCAGCGAGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGCTCGCCAATAAAGGCGCACTCCAGGAGCTCTTCGGGGTGAGACACACCGCTCAGCGAGAG 660

Qy    661 AAGACGGAGTGGCATGGGAGGTGGCCTTCGTACCTGGTCAGTTATGGGGGAGCAGCGGGGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAGACGGAGTGGCATGGGAGGTGGCCTTCGTACCTGGTCAGTTATGGGGGAGCAGCGGGGCC 720

Qy    721 TGGGACCTAGCCCTCCATACCTGGGAGCAGAGGGGGCTGAGGTCACTGTGCGCCCAAGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGGACCTAGCCCTCCATACCTGGGAGCAGAGGGGGCTGAGGTCACTGTGCGCCCAAGCC 780

Qy    781 CAGGAAGGGGCAGGCCACTCCCCCTCCATCCCCTACAGCCCAAGCCAGTGCCAAGCC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAGGAAGGGGCAGGCCACTCCCCCTCCATCCCCTACAGCCCAAGCCAGTGCCAAGCC 840

Qy    841 TCTCCCCAGCCAACCCCACCTCCACCGCAGTGCTAATGCCCTGGATCCCATGAATTGCCGGCG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTCCCCAGCCAACCCCACCTCCACCGCAGTGCTAATGCCCTGGATCCCATGAATTGCCGGCG 900

Qy    901 GGGGTGCACCCAGGGGCTCAGAGAGAAGGGGTTTGAGACAGTGCCTGACACATCTGGACGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGGTGCACCCAGGGGCTCAGAGAGAAGGGGTTTGAGACAGTGCCTGACACATCTGGACGC 960

Qy    961 CGCTGGAGAGAAAATCTCTGCCTCACTCTTCCAAGCTCTTCCAAAGCTCCCCAGACCAT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGCTGGAGAGAAAATCTCTGCCTCACTCTTCCAAGCTCTTCCAAAGCTCCCCAGACCAT 1020

Qy   1021 GAGTCTCAAGGCCAGGAGTCACCCAAGCGCCCCCACATCCACACAGCAGTGCTGGGGAGCTGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAGTCTCAAGGCCAGGAGTCACCCAAGCGCCCCCACATCCACACAGCAGTGCTGGGGAGCTGG 1080

Qy   1081 GGATCCCACCTCAGCCCTAGCCACCCAGAGAGCAGGAGGCTCCTGGGGACCAAATGG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGATCCCACCTCAGCCCTAGCCACCCAGAGAGCAGGAGGCTCCTGGGGACCAAATGG 1140

Qy   1141 CCTCTGGATGAAAAGGTCAGGAAATTTACTACACAGAAAATCAGAGAAAGAGAGAGAGAGAAA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCTCTGGATGAAAAGGTCAGGAAATTTACTACACAGAAAATCAGAGAAAGAGAGAGAGAGAAA 1200

Qy   1201 TCAGAGAAAGGCAGGCCCCCCATGGGCAGCGGGTGGTAGGGAACGCCCCCCACAGGCGGCACACC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAGAAAGGCAGGCCCCCCATGGGCAGCGGGTGGTAGGGAACGCCCCCCACAGGCGGCACACC 1260
```

```
Qy   1261 AGCCTACAGCCCACCCACCACCCATGGGAGGAGCCTTCTGTGAGAGAGAGAGAGCTTCGTTTCCACA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGCCTACAGCCCACCCACCACCCATGGGAGGAGCCTTCTGTGAGAGAGAGAGAGCTTCGTTTCCACA 1320

Qy   1321 TGGCCCTTGGAAAAATGAGGATTTTAACCAAAAAAATTCACACAGCTGCTACTTCTACAAAGA 1380
Db   1321 TGGCCCTGGAAAAATGAGGATTTTAACCAAAAAAATTCACACAGCTGCTACTTCTACAAAGA 1380

Qy   1381 CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGGAAGTGGGCTGATTATGTGGAGGAG 1440
Db   1381 CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGGAAGTGGGCTGATTATGTGGAGGAG 1440

Qy   1441 AATCGAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGGACTACCCAAGAA 1500
Db   1441 AATCGAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGGACTACCCAAGAA 1500

Qy   1501 CCTCGCCATAGTCATACTTGCAGGGGGGCTGGAAGTGGGGAGTCAACACTGGCCAGGCAG 1560
Db   1501 CCTCGCCATAGTCATACTTGCAGGGGGGCTGGAAGTGGGGAGTCAACACTGGCCAGGCAG 1560

Qy   1561 GTGAAGGAAGCCTGGGGAGGAGGCCAGGTTCCCAGAGGAGGGCCAGTGCATGTGAGAA 1620
Db   1561 GTGAAGGAAGCCTGGGGAGGAGGCCAGGTTCCCAGAGGAGGGCCAGTGCATGTGAGAA 1620

Qy   1621 TTCAGCTGCAGAGAGCTGGCCCAGGTGGTGAGGTCTCCGCTGAGGTCATCGGGAAAA 1680
Db   1621 TTCAGCTGCAGAGAGCTGGCCCAGGTGGTGAGGTCTCCGCTGAGGTCATCGGGAAAA 1680

Qy   1681 GATGGGACAGGCCACTCCGGCTCCCATTAGACAGATCCTATGGCCAGAGGGGCTGCTC 1740
Db   1681 GATGGGACAGGCCACTCCGGCTCCCATTAGACAGATCCTATGGCCAGAGGGGCTGCTC 1740

Qy   1741 TTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTTCTTGCAGAGTTCTGAGCTC 1800
Db   1741 TTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTTCTTGCAGAGTTCTGAGCTC 1800

Qy   1801 TGTCTGCACTGGAGCCCACAGCCGCGGATGCACTGGCTGGGAATGGGTTTGGGGAAA 1860
Db   1801 TGTCTGCACTGGAGCCCACAGCCGCGGATGCACTGGCTGGGAATGGGTTTGGGGAAA 1860

Qy   1861 ACTATACTTCCCGAGGCAGCCATCCTTCCTGATCACGGGCCACACGTTGGGTTAGAGCTC 1920
Db   1861 ACTATACTTCCCGAGGCAGCCATCCTTCCTGATCACGGGCCACACGTTGGGTTAGAGCTC 1920

Qy   1921 ATTCCTTCTTTGGAGCAGGCACGTTGGGTTGGGGTTCCTGGGGTTCTCTGAGTCCAGCAGG 1980
Db   1921 ATTCCTTCTTTGGAGCAGGCACGTTGGGTTGGGGTTCCTGGGGTTCTCTGAGTCCAGCAGG 1980

Qy   1981 AAGGAAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCAATTAGACCTTTAGGTTG 2040
Db   1981 AAGGAAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCAATTAGACCTTTAGGTTG 2040

Qy   2041 GTCAAATCAAACAAAGAGAGCTCTGGGCCCCTGTGTTCTTGTTGCTTGTCTTGGCTGGCC 2100
Db   2041 GTCAAATCAAACAAAGAGAGCTCTGGGCCCCTGTGTTCTTGTTGCTTGTCTTGGCTGGCC 2100

Qy   2101 TGCACTTGCCTGATGCAGCAGATGGAAGCGGAAGAAAAACTCACACTGACTTCCAAGACC 2160
Db   2101 TGCACTTGCCTGATGCAGCAGATGGAAGCGGAAGAAAAACTCACACTGACTTCCAAGACC 2160

Qy   2161 ACCACAACCCTCTGTCTCTACATTACCTTGCCCAAGGCTCAGCCATTGGGACCC 2220
Db   2161 ACCACAACCCTCTGTCTCTACATTACCTTGCCCAAGGCTCAGCCATTGGGACCC 2220

Qy   2221 CAGCTCAGAGACCTTCTGCTCTCTGGCTGCTGAGGGCATCTGGCAAAAAAAGACCCTTTTC 2280
Db   2221 CAGCTCAGAGACCTTCTGCTCTCTGGCTGCTGAGGGCATCTGGCAAAAAAAGACCCTTTTC 2280

Qy   2281 AGTCCAGATGACCTCAGGAGCATGGGGTTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
Db   2281 AGTCCAGATGACCTCAGGAGCATGGGGTTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340

Qy   2341 ATGGGGTATTCTTCAAGAGCACCCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC 2400
Db   2341 ATGGGGTATTCTTCAAGAGCACCCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC 2400
```

```
Db   2341 ATGGGTATTCTTCAAGAGCACCCCCATCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2401 CAAGAGTTCTTTTGCAGCAATGTCCTTATGTCTTCAGGTCAGGATGAGAAGGGGGAGAAACAT 2460

Db   2401 CAAGAGTTCTTTTGCAGCAATGTCCTTATGTCTTCAGGTCAGGATGAGAAGGGGGAGAAACAT 2460

Qy   2461 TCTAATTGCATCATAGATTTTGGAAAAGACGGCTAGAAGCATATATGGAATACATGGCCTGTTT 2520
Db   2461 TCTAATTGCATCATAGATTTTGGAAAAGACGGCTAGAAGCATATATGGAATACATGGCCTGTTT 2520

Qy   2521 GGGGCCATCAAACCACACGTTTCCTATTGGGCCCTGTGTTAAGTGATGATGAGGGGAGGATG 2580
Db   2521 GGGGCCATCAAACCACACGTTTCCTATTGGGCCCTGTGTTAAGTGATGATGAGGGGAGGATG 2580

Qy   2581 GAGAACATCTTTCACTGCCGGCTGCTGTCTCCAGGGGGAGGGAACCTGAATGCAGTGGGTCCC 2640
Db   2581 GAGAACATCTTTCACTGCCGGCTGCTGTCTCCAGGGGGAGGGAACCTGAATGCAGTGGGTCCC 2640

Qy   2641 CTGCAGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCTCCACTGCTTGTTGTACGGTGGTCCTG 2700
Db   2641 CTGCAGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCTCCACTGCTTGTTGTACGGTGGTCCTG 2700

Qy   2701 AACAAAAACGTTCCTGACACAAAGTGATGGGCCCATTTCGAGAAATGGGCATGTGGTTAGAA 2760
Db   2701 AACAAAAACGTTCCTGACACAAAGTGATGGGCCCATTTCGAGAAATGGGCATGTGGTTAGAA 2760

Qy   2761 ACAGACATGGAGCTCTTAGTGTTGCACTTTCTGCATTAAATTCAGCGCGCCACGTGAAGAG 2820
Db   2761 ACAGACATGGAGCTCTTAGTGTTGCACTTTCTGCATTAAATTCAGCGCGCCACGTGAAGAG 2820

Qy   2821 CTTCAGCTGATTGAGGGCAGGCAGCACAGATCAACATGAGGTAGTCCTG 2880
Db   2821 CTTCAGCTGATTGAGGGCAGGCAGCACAGATCAACATGAGGTAGTCCTG 2880

Qy   2881 TTCAGGTGGGTCCCAGTCACAGATGCCATTGGGCAGATTCTTCTCCGGTCCTCAAGGTC 2940
Db   2881 TTCAGGTGGGTCCCAGTCACAGATGCCATTGGGCAGATTCTTCTCCGGTCCTCAAGGTC 2940

Qy   2941 ACCAGAAACCTGAAGGAGTGGACCTTAAGTGGAAAACTCGGTCAGCCACTCTGCAGTGAAG 3000
Db   2941 ACCAGAAACCTGAAGGAGTGGACCTTAAGTGGAAAACTCGGTCAGCCACTCTGCAGTGAAG 3000

Qy   3001 AGTCCTTGTTAAGGACCCCTGAGACGGCCCTGCCTCCGAGACCCTTGGCTGGC 3060
Db   3001 AGTCCTTGTTAAGGACCCCTGAGACGGCCCTGCCTCCGAGACCCTTGGCTGGC 3060

Qy   3061 TGTTGGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGGCCTTGAGAGCCAACCAGACC 3120
Db   3061 TGTTGGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGGCCTTGAGAGCCAACCAGACC 3120

Qy   3121 CTGACCGAGCTGGACCTGAGCTTCAAATGTGCTCACGGACTCTGAGGAGCCAAACACCTTTGC 3180
Db   3121 CTGACCGAGCTGGACCTGAGCTTCAAATGTGCTCACGGACTCTGAGGAGCCAAACACCTTTGC 3180

Qy   3181 CAGAGGACTGAGACAGCCGAGGCTGCAAGCTACAGCGACTGGTCAGCTGTTCAGCTC 3240
Db   3181 CAGAGGACTGAGACAGCCGAGGCTGCAAGCTACAGCGACTGGTCAGCTGTTCAGCTC 3240

Qy   3241 ACGTCTGACTGCTGCCAGGACCTGGCCTGCCTTGAGGTGCCAGCCTGGAAGGAG 3300
Db   3241 ACGTCTGACTGCTGCCAGGACCTGGCCTGCCTTGAGGTGCCAGCCTGGAAGGAG 3300

Qy   3301 CTAGACCTGCAGCGAGAACAAACCTGGACGTTGGCGTGGCACTGCTTCTGTTGAGGGGGCTC 3360
Db   3301 CTAGACCTGCAGCGAGAACAAACCTGGACGTTGGCGTGGCACTGCTTCTGTTGAGGGGGCTC 3360

Qy   3361 AGGCATCCTGCCTGCAAACTCATACGCCTGGGGCTGGACAACTCTGAGTGATGAG 3420
Db   3361 AGGCATCCTGCCTGCAAACTCATACGCCTGGGGCTGGACAACTCTGAGTGATGAG 3420

Qy   3421 ATGAGGCAGGAACTGAGGGCCTGGAGCAGGGAGAAAACCTCAGCTGCTCATCTTCAGCAGA 3480
Db   3421 ATGAGGCAGGAACTGAGGGCCTGGAGCAGGGAGAAAACCTCAGCTGCTCATCTTCAGCAGA 3480
```

Qy   3421 ATGAGGCAGGAACTGAGGGGCCCTGAGCGAGAGAAACCTGAGCTGCTCAATCTTCAGCAGA 3480

Db   3481 CGGGAAACCAAGTGTGATGACCCCTACTGAGGGCCTTGGATACGGAGGAGAGGAATAATAGC 3540

Db   3481 CGGGAAACCAAGTGTGATGACCCCTACTGAGGGCCTTGGATACGGAGGAGAATGAGTAATAGC 3540

Qy   3541 ACATCCTCACTCAAGCGGCACAGAGACTCGGATCAGAGGAGGCGCGGCTTCCCATGTTGCTCAG 3600

Db   3541 ACATCCTCACTCAAGCGGCACAGAGACTCGGATCAGAGGAGGCGCGGCTTCCCATGTTGCTCAG 3600

Qy   3601 GCTAATTCTCAAAACTCCTGACGTGAGCAGCAGATCTTCCCCAATTGCTCAGAGATTGCAGAGGAA 3660

Db   3601 GCTAATTCTCAAAACTCCTGACGTGAGCAGCAGATCTTCCCCAATTGCTCAGAGATTGCAGAGGAA 3660

Qy   3661 AGCTCCCCAGAGAGGTAGTTACCGGTTGGGAACTTCTTGTGTCGGTGGCCTTCTCCTGGCCTTCTCAAGGG 3720

Db   3661 AGCTCCCCAGAGAGGTAGTTACCGGTTGGGAACTTCTTGTGTCGGTGGCCTTCTCCTGGCCTTCTCAAGGG 3720

Qy   3721 GACCTGCAATACGGAGGACCTTTGGGGACTGACGATGACTTTCGGGGGCCCCAACGGGGGCCCTGTG 3780

Db   3721 GACCTGCAATACGGAGGACCTTTGGGGACTGACGATGACTTTCGGGGGCCCCAACGGGGGCCCTGTG 3780

Qy   3781 GCTACTGAGGTAGTTGACAAAGAAAAGAAGAACTTGTTCACTTCCCCGTRAGCTGGC 3840

Db   3781 GCTACTGAGGTAGTTGACAAAGAAAAGAAGAACTTGTTCACTTCCCTGTAGCTGGC 3840

Qy   3841 TCCTACCGGTTGGCCCAACACGGGGTCTCTGCTTTGTGATGAGAGAAGCGGTGACCGTTGAG 3900

Db   3841 TCCTACCGGTTGGCCCAACACGGGGTCTCTGCTTTGTGATGAGAGAAGCGGTGACCGTTGAG 3900

Qy   3901 ATTGAATTCTGTGTGTGGGACCAGTTCCTGGGTGAGATCAAACCCACAGCACAGTCGGATG 3960

Db   3901 ATTGAATTCTGTGTGTGGGACCAGTTCCTGGGTGAGATCAAACCCACAGCACAGTCGGATG 3960

Qy   3961 GTGGCAGGGCGCCTCTGCTGGACATCAAGGCTGAGCGTGTGGAAGCTGTGCGCACCTC 4020

Db   3961 GTGGCAGGGCGCCTCTGCTGGACATCAAGGCTGAGCGTGTGGAAGCTGTGCGCACCTC 4020

Qy   4021 CCTCACTTTGTGGGCTCTCCAAGGGGGCCCATGTGACACATCCTGTTCCAAATGGCCCAC 4080

Db   4021 CCTCACTTTGTGGGCTCTCCAAGGGGGCCCATGTGACACATCCTGTTCCAAATGGCCCAC 4080

Qy   4081 TTTAAAGAGGAGGGGGATGCTCCCAGGTCCTGGAGGTCCTGGAGCTGCATCACATAGTT 4140

Db   4081 TTTAAAGAGGAGGGGGATGCTCCCAGGTCCTGGAGGTCCTGGAGCTGCATCACATAGTT 4140

Qy   4141 CTGGAAAACCCCAGGTCTTCTCCCCCCTTGGGAGTCCCCTGAAAATGATCCATAAATGCCCTG 4200

Db   4141 CTGGAAAACCCCAGGTCTTCTCCCCCTTGGGAGTCCCCTGAAAATGATCCATAAATGCCCTG 4200

Qy   4201 CGCTTCATTCCCGTGTCACCTCGTGGTGTTGCTTTACCACGGCGTCCATCCTGAGGAAGTC 4260

Db   4201 CGCTTCATTCCCGTGTCACCTCGTGGTGTTGCTTTACCACGGCGTCCATCCTGAGGAAGTC 4260

Qy   4261 ACCTTCCACCTCTACCTGATCCCAAGTGACTGACTGCTCCATTCGGGAGGAACTGGAGCTCTGC 4320

Db   4261 ACCTTCCACCTCTACCTGATCCCAAGTGACTGACTGCTCCATTCGGGAGGAACTGGAGCTCTGC 4320

Qy   4321 TATCGAAGCCCCTGGAGAAGACCAGCTGTTCTCGGAGTTCTACGTTGGCCACTTGGGATCA 4380

Db   4321 TATCGAAGCCCCTGGAGAAGACCAGCTGTTCTCGGAGTTCTACGTTGGCCACTTGGGATCA 4380

Qy   4381 GGGATCAGGCTGCAAGTGAAAGATGAAAGATAGACTTCTGGTGTGTGGGAGGCCTTGGTG 4440

Db   4381 GGGATCAGGCTGCAAGTGAAAGATGAAAGATAGACTTCTGGTGTGTGGGAGGGCCTTGGTG 4440

Qy   4441 AAACCAGGAGAGACTCTCATGCCTGCAACTACTCTCGATCCCTCCAGCCCGCATAGCCGTACCT 4500

Db   4441 AAACCAGGAGAGACTCTCATGCCTGCAACTACTCTCGATCCCTCCAGCCCGCATAGCCGTACCT 4500

Qy   4501 TCACCTCTGGATGCCCCGCAGTTGCTCGACTTTGTGGGACCAGTATCGGAGCAGCAGCTGATA 4560

Db   4501 TCACCTCTGGATGCCCCGCAGTTGCTCACTTTGCTGGGACCAGTATCGGAGCAGCAGCTGATA 4560

Qy   4561 GCCCGGAGTGACATCGGTGGAGGTTGTTGCTTGGACAAAACTGCATGGAACAGGTGCTGCTGAGCCAG 4620

Db   4561 GCCCGGAGTGACATCGGTGGAGGTTGTTGCTTGGACAAAACTGCATGGAACAGGTGCTGCTGAGCCAG 4620

Qy   4621 GAGCAGTACGGAGAGGGGTGCTGGCTGAGAACGAGAGGCCCAGAGATGCGGAAGCTGTTC 4680

Db   4621 GAGCAGTACGGAGAGGGGTGCTGGCTGAGAACGAGAGGCCCAGAGATGCGGAAGCTGTTC 4680

Qy   4681 AGCTTGAGCCAGTCCTGGAGCACCGGAAGTGCAAAGATGGACTCTACCAAGCCCTGAAGGAG 4740

Db   4681 AGCTTGAGCCAGTCCTGGAGCACCGGAAGTGCAAAGATGGACTCTACCAAGCCCTGAAGGAG 4740

Qy   4741 ACCCATCCTCACCCTCACTCATTATGGAACTCTCGGGAGGAGGCAAAAAGGGACTCCTGCCA 4800

Db   4741 ACCCATCCTCACCCTCACTCATTATGGAACTCTCGGGAGGAGGCAAAAAGGGACTCCTGCCA 4800

Qy   4801 CTCAGCAGTGAAGTATCAACACAGCCCCTTGACCCTTGAGTCCTGGCTTTGGGCTGACCC 4860

Db   4801 CTCAGCAGTGAAGTATCAACACAGCCCCTTGACCCTTGAGTCCTGGCTTTGGGCTGACCC 4860

Qy   4861 TTCTTTGGGTCTCAGTTTCTTTCTTCTCTGCAAACAAGTTGCATCTCGGTTTGCCTTCCAGCA 4920

Db   4861 TTCTTTGGGTCTCAGTTTCTTTCTTCTCTGCAAACAAGTTGCATCTCGGTTTGCCTTCCAGCA 4920

Qy   4921 CTAAAGTAATGGAACTTTGATGATGCCCTTGGCATTATGTTGTCCATGCCAGGGATG 4980

Db   4921 CTAAAGTAATGGAACTTTGATGATGCCCTTGGCATTATGTTGTCCATGCCAGGGATG 4980

Qy   4981 CCACAGGGGGCCCCAGTCCAGGTGGCCTAAACAGCATCCAGGGGAATGTCCATCTCGGAGCT 5040

Db   4981 CCACAGGGGGCCCCAGTCCAGGTGGCCTAAACAGCATCCAGGGGAATGTCCATCTCGGAGCT 5040

Qy   5041 GGCAAGACCCCCTGCAGACCCTCATCCAGGCGCCAAAAAATGGAGCAGCAGCCAAGCCTAGA 5100

Db   5041 GGCAAGACCCCCTGCAGACCCTCATCCAGGCGCCAAAAAATGGAGCAGCAGCCAAGCCTAGA 5100

Qy   5101 TGGGCTGTGTCACAGGGGTGAGCCCCAAAAATGGGGTTCAGCGTGGAGGACAATGGAACCATT 5160

Db   5101 TGGGCTGTGTCACAGGGGTGAGCCCCAAAAATGGGGTTCAGCGTGGAGGACAATGGAACCATT 5160

Qy   5161 TTGGCTTTTGTTACAGGAGATCACAGAGCCAACAGAGTGAAAGTGGAAGGAAGTTT 5220

Db   5161 TTGGCTTTTGTTACAGGAGATCACAGAGCCAACAGAGTGAAAGTGGAAGGAAGTTT 5220

Qy   5221 ATTCAGAGAAATAAAGGGAGTATCACAGCTCCTTTTAGAATTGTCTAGCAGGCTTTCCAGTT 5280

Db   5221 ATTCAGAGAAATAAAGGGAGTATCACAGCTCCTTTTAGAATTGTCTAGCAGGCTTTCCAGTT 5280

Qy   5281 TTTACCAGAAAAACCCTATAAAATTTAAAATTTTTTACTTAAATTAAGAATTAAAAAAT 5340

Db   5281 TTTACCAGAAAAACCCTATAAAATTTAAAATTTTTTACTTAAATTAAGAATTAAAAAAT 5340

Qy   5401 ACAAAAAAGAAAAAATGAAAAATAAAGGGAATTAAGAAGTTACCTAC 5444

Db   5401 ACAAAAAAGAAAAAATGAAAAATAAAGGGAATTAAGAAGTTACCTAC 5444

```
REFERENCE   1  (bases 1 to 5404).
AUTHORS     Martinon,F., Hofmann,K. and Tschopp,J.
TITLE       The Pyrin domain: a possible member of the death domain-fold family
            implicated in apoptosis and inflammation
JOURNAL     Curr. Biol. 11 (4), R118-R120 (2001)
MEDLINE     11148093
PUBMED      11250163
REFERENCE   2  (bases 1 to 5404)
AUTHORS     Martinon,F., Hofmann,K. and Tschopp,J.
TITLE       Direct Submission
JOURNAL     Submitted (28-SEP-2000) Institute of Biochemistry, University of
            Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
FEATURES             Location/Qualifiers
     source          1..5404
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             519..4808
                     /note="contains NACHT, LRR and PYD domains"
                     /codon_start=1
                     /product="NALP1"
                     /protein_id="AAG30288.1"
                     /db_xref="GI:11096303"
                     /translation="MAGGAWGRLACYLFELKKEELKEPQLLLANKAHSRSSSGETPAQ
                     PEKYSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQAQEGAGHSPSFPYSPSE
                     PHLGSPSQPFSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQAL
                     PSSPDHESPSQEEPNAPTSTAVLGSWGSPPQPSLAPREQEAPQTGWPLDETSGITYYTE
                     IRREREKSEKGRRPWAAVVGTPPQAHTSLQPHHHPWEESVRESLCSTWPWKNEDFNQG
                     KFTQLLLLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFGPGLDTQEPRIVLQG
                     AAIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALGSILGKTILP
                     EASFLITARTTALQNLIPSLEQARWEVLGFSESSRKEYFYRYFTDERQAIRAPRLVK
                     SNKELWALCLVPWWSWLACTCLMQQMKRREKLITLSKTTTLCLHYLAQALQAQPLGP
                     QLRDLCSLAAEGIWQKKKLFSPDDLRKHGLDGAIISTFLPKMGILQEHPIPLSYSFIHL
                     CFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAYGIHGLFGASTTRFLLGLLSDEG
                     EREMENIFHCRLSQGRNLMQMVPSLQLLLQPHSLESLHCLYETRNKTFLTQVMAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKKLQLIEBGRQHRSTWSPTMVVLFRNVPVTDAYWQ
                     ILFSVLKVTRNLKELDLSGNSLSHSAAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKD
                     LAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDL
                     ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELR
                     ALEQEKPQLLIFSRKPSWMTPIEGLDTGEMSNSTSLRQRLGSSERAASHVAQANLK
                     LLDVSKIPIAEIAEESSPEVPVELLCVPSPASQGDLHTKPLGTDDDRWGPGPVAT
                     EVDKEKNLYRVHPPVAGSYRWNTGLCPVMREAVTVEIBFCVWDQFLGEINPQHSWM
                     VAGPLDDIRAEPGAVEAVRLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARVELHH
                     IVLENPSFSPLGVLLKMIHNALRFYPIVTSVVLLYHRVHPEEVTFHLYILPSDCSIRKE
                     LELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP
                     ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVLDKLHGQVLSQEQYERVLAENTR
                     PSQMRKLFSLSQSWDRKCKDGLIYQALKSTHPHLIMELWEKGSKGLLPLSS"
BASE COUNT      1284 a   1518 c   1467 g   1135 t
ORIGIN
```

```
Query Match             99.2%;  Score 5400.8;  DB 9;  Length 5404;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches  2;
Matches 5402; Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy       5 CAGGGGCTGGAGAGGTCTGAAGAAAACCTGGGAGACCAGCAGCGGGGGGCTCCACTCTGGGT    64
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAGGGGCCTCGAGAGGTCTGAAGAAAACCTGGGAGACCAGCAGCGGGGGGCTCCACTCTGGGT    60

Qy      65 TCTGAAAGCCCATTCCCTGCTTCCTGCGGCTCCCAGCCCACCCTCTTCTTCTTCAGCCTTGCAG   124
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCTGAAAGCCCATTCCCTGCTTCCTGCGGCTCCCAGCCCACCCTCTTCTTCTTCAGCCTTGCAG   120

Qy     125 CTCAAGAGGGTTGGATCTCAGGGACCCAGGACCCAGGGGGAGGAATCTGAGGAGACACAGA   184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTCAAGGGGTTGGATCTCAGGGAGTCCAGGACCCAGGGGGGAGAGAATCTGAGGAGAACACAGA   180

Qy     185 ACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCACCATCTCCCCTCACCCTCACCCTC   244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCACCATCTCCCCTCACCCTCACCCTC   240

Qy     245 CCTGCCTGGGCCCTGGACCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGTGTCT   304
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTGCCTGGGCCCTGGACCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGTGTCT   300
```

```
Qy     305 TGCAGGGCCCCCTCTGGGCTTCCTCCCTCCCCTCCCCTGGCTTTTCCTACCACTCCCCTCTATCGGC   364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGCAGGGCCCCCTCTGGGCTTCCTCCCTCCCCTCCCCTGGCTTTTCCTACCACTCCCCTCTATCGGC   360

Qy     365 GTCTATCTGTAGGTGCCCTGGGGATTTTATAAAAACTGGGTTCCGAAATGCTGAATAAGAGACG   424
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTCTATCTGTAGGTGCCCTGGGGATTTTATAAAAACTGGGTTCCGAAATGCTGAATAAGAGACG   420

Qy     425 GTAAGAGCCAAGGCAAAGGACAGCACTGTTTCTCTGCCTGGATACCCTCACCCTGG   484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTAAGAGCCAAGGCAAAGGACAGCACTGTTTCTCTGCCTGGATACCCTCACCCTGG   480

Qy     485 GAAACATCCCCAGACACACCCTCTTTAACTCCGGACACAGAGTGGCTGCGGGACGCCTGGGGCC   544
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GAACATCCCCAGACACCCTCTTTAACTCCGGACACAGAGTGGCTGCGGGACGCCTGGGGCC   540

Qy     545 GCCTGGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAGCTGAAGGAGGTTCCAGCTTCTGC   604
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GCCTGGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAGCTGAAGGAGGTTCCAGCTTCTGC   600

Qy     605 TCGCCAATAAACGCACTCCAGGAGCTCTTCGGGGTGAGACACCCGCTCAGCCAGGAGAAGA   664
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCGCCAATAAACGCACTCCAGGAGCTCTTCGGGGTGAGACACCCGCTCAGCCAGGAGAAGA   660

Qy     665 CGAGTGGCATGGAGGTGGGCCTTCGTACCTGTCGTGGGCTCAGTAATGGGGGAGCAGGGGCCTGGG   724
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGAGTGGCATGGAGGTGGGCCTTCGTACCTGTCGTGGGCTCAGTAATGGGGGGAGCAGGGGCCTGGG   720

Qy     725 ACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTTGCGGCCCAAGCCCAGG   784
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACCTAGCCCTCCATACCTGGGAGCAGATGGGGGCTGAGGTCACTGTTGCGGCCCAAGCCCAGG   780

Qy     785 AAGGGGCGAGGCCACTCTCCCTCATTCCCCTACAGCCCAAGTGAACCCCACCTGGGGGTCTC   844
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AAGGGGCGAGGCCACTCTCCCTCATTCCCCTACAGCCCAAGTGAACCCCACCTGGGGGTCTC   840

Qy     845 CCAGCCAAACCCACCTCCAGCCAGTTGCTAATGCCCTGGGATCACATGAATTGCGGGCGGGGT   904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCAGCCAAACCCACCTCCAGCCAGTTGCTAATCCCTGGGATCACATGAATTGCGGGCGGGGT   900

Qy     905 GCACCCCAGGGGTCAGGAGGAGAGGAGGTTTTGAGACAGCTGGCCTGACACACATCTGGACGCGCGT   964
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GCACCCCAGGGGTCAGGAGGAGAGGAGGTTTTGAGACAGCTGGCCTGACACACATCTGGACGCGCT   960

Qy     965 GGAGAGAAATCTGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCCAGACCATGAGT   1024
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GGAGAGAAATCTGCCTCACTCCTCTACCAAGCTCTTCCAAGCTCCCCAGACCATGAGT   1020

Qy    1025 CTCCAAGCCAGGAGTGTCACCCAACGCCCCCACATCCACAGCAGTGTCTGGGGAGCTGGGGT   1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CTCCAAGCCAGGAGTCACCCAACGCCCCCACATCCACAGCAGTGCTGGGGAGCTGGGGAT   1080

Qy    1085 CCCCACCTCAGCCCAGCCTAGCACCCAGGACCAGAGAGCAGCAGGAGGGCTCCTGGGAAATGGCCTC   1144
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CCCCACCTCAGCCCAGCCTAGCACCCAGGACCAGAGAGCAGCAGGAGGGCTCCTGGGAAATGGCCTC   1140

Qy    1145 TGGATGAAAACGTCAGGAGAATTACTACACAGAAATCAGAGAGAGAGAGAGAGAGAAATCAG   1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TGGATGAAAACGTCAGGAGAATTTACTACACAGAAATCAGAGAGAGAGAGAGAGAGAAATCAG   1200

Qy    1205 AGAAACGCAGGCGCCCCCCATGGGCAGCGCGGTGGTAGGAACGGGCCCCACAGGGGCACACCAGCC   1264
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 AGAAAGGCAGGCGCCCCCCATGGGCAGCGCGGTGGTAGGAACGGGCCCCACAGGGGCACACCAGCC   1260

Qy    1265 TACAGCCCCACCACCACCCCATGGGGGAGCCTTCGTTGTGAGAGGACGGGCCCTCGTTCCACATGGC   1324
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TACAGCCCCACCACCACCCCATGGGGGAGCCTTCGTTGTGAGAGGACGGGCCCTCGTTCCACATGGC   1320

Qy    1325 CCTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAGCTGCTACTTCTACAAAAGACCTC   1384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCTGGAAAAATGAGGATTTTAACCAAAAAATTCACACAGCTGCTACTTCTACAAAAGACCTC   1380
```

```
Qy   1385 ACCCCAGGAGGCCAAGATCCCCTGGTCAGAGAGAACTGGCCTGATTAGTAGTGGGAGGAATC 1444
Db   1381 ACCCCAGGAGGCCAAGATCCCCTGGTCAGAGAGAACTGGCCTGATTAGTAGTGGGAGGAATC 1440

Qy   1445 GAGGACACATTTAATTGAGATCAGAGATCAGAGACTTATTTGCCCACGGCCTGATTAGTGGGAGGAAGCTTC 1504
Db   1441 GAGGACACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGAAACCTC 1500

Qy   1505 GCATAGTCATACTGCAGGGGGGCTTCGGAATTGGGGAGTCAACACTGGCCCAGGCCAGGTGA 1564
Db   1501 GCATAGTCATACTGCAGGGGGCTGCTGAGAATTGGGGAGTCAACACTGGCCCAGGCCAGGTGA 1560

Qy   1565 AGGAAGCCTGGGGGAGAGGCCAGGCTGTATGGGGAGACCGGCTTCCAGCCATGTCTTCTACTTCA 1624
Db   1561 AGGAAGCCTGGGGGAGGACGGCCAGCTGTATGGGGGACCGGCTTCCAGCCATGTCTTCTACTTCA 1620

Qy   1625 GCTGCAGAGAGCTGGGCCCCAGTCCAAGGTGGTGTTGAGTCTTCGCTGCTGCTGGAAAAGAATG 1684
Db   1621 GCTGCAGAGAGCTGGCCCCAGTCCAAGGTGGTGAGTCTTCGCTGCTGCTGGAAAAGATG 1680

Qy   1685 GGACAGCAGCCACTCCGGCTTCCCATTAGACAGATCCTGGTCCTAGAGGCCAGAGGCGGCTCGCTCTTCA 1744
Db   1681 GGACAGCGCACTCCGGGCTCCCATTAGACAGATCCTGGTCCTAGAGGCCAGAGGCGGCTGCTCTTCA 1740

Qy   1745 TCCTTCCCGATGGTGTAGATGAGCCCAGGATGGGGTTCTTGCACGAGAGCCCAGTTCTTGAGCTCTGTC 1804
Db   1741 TCCTTCGATGGTGTAGATGAGCCCAGGATGGGGGTTCTTGCACGAGAGCCCAGTTCTTGAGCTCTGTC 1800

Qy   1805 TGCACTGGAGCCAGCCACCACAGCCGGGATGCACTCGGCCAGTTGCTGGGAGAAAACTA 1864
Db   1801 TGCACTGGAGCCAGCCACCACAGCCGGGATGCACTCGGCCAGTTGCTGGGAGAAAACTA 1860

Qy   1865 TACTTCCCGAGGGCATCCTTCCTGATCACGGCTCGGACACCAGCTCTGCGAGAACCTCATTC 1924
Db   1861 TACTTCCCGAGGGCATCCTTCCTGATCACGGCTCGGACACCAGCTCTGCGAGAACCTCATTC 1920

Qy   1925 CTTCTTTGGAGCAGGCACCGTTGGGGTAGGTCTTCTGAGTCCAGCAGCGAAGG 1984
Db   1921 CTTCTTTGGAGCAGGCACCGTTGGGGTAGGTCTTCTGAGTCCAGCAGCGAAGG 1980

Qy   1985 AATATTTCTACAGATATTTCACAGATGAAGAGGGCAAGCAAATTAGGTGGGTCA 2044
Db   1981 AATATTTCTACAGATATTTCACAGATGAAGAGGGCAAGCAAATTAGGTGGGTCA 2040

Qy   2045 AATCAAAACAAAAGACTCTGGGGGCCCCTGTGTGTCTTGTGGCCCTGGGCTGGCCTGCA 2104
Db   2041 AATCAAAACAAAAGACTCTGGGGGCCCCTGTGTCTTGTGGCCCTGGGCTGGCCTGCA 2100

Qy   2105 CTTGCCTGATGCAGCGATGAAGGGGAAGGGGAAGAAAACTCACACTGACTTCCAAGACCACCA 2164
Db   2101 CTTGCCTGATGCAGCGATGAAGGGGAAGAAAACTCACACTGACTTCCAAGACCACCA 2160

Qy   2165 CAACCCTCGTCTACATTACCTTGCCCAGGCTCTCCCAGCCTCAGCCATTGGGACCCCAGC 2224
Db   2161 CAACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCCAGCCTCAGCCATTGGGACCCCAGC 2220

Qy   2225 TCAGAGACCTCTCCTCTCTGGCTGCTGAGGGCATCCTGGGCAAAAAAGACCCTTTTCAGTC 2284
Db   2221 TCAGAGACCTCTCCTCTCTGGCTGCTGAGGGCATCGGCAAAAAAGACCCTTTTCAGTC 2280

Qy   2285 CAGATGACCTCAGGAAGCATGGGGTTAGATGGGGCCCATCCATCCACCTTCTTGAAGATGG 2344
Db   2281 CAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCCATCTCCACCTTCTTGAAGATGG 2340

Qy   2345 GTAATTCTTCAAAGAGCACCCCATCCCTCTGAGCTACACTCACCTCTGTTTCCAAG 2404
Db   2341 GTATTCTTCAAAGAGCACCCCATCCCTCTGAGCTACACCTCATCTTCAGCAGACGGA 2400

Qy   2405 AGTTCTTTTGCAGCAATGTCCTATGTCTTGGAGGATGAGAGGGGGAGGGGGTAAACATTCTA 2464
Db   2401 AGTTCTTTGCAGCAATGTCCTAGTCTCTTGGAGGATGAGAGGGGAGGGGGTAAACATTCTA 2460
```

```
Qy   2465 ATTGCATCAATAGATTTGGAAAAGACGCTTAGAAGCAATATGGGAATACATGGGCCTGTTTGGGG 2524
Db   2461 ATTGCATCAATAGATTTGAAAAAAGACGCTTAGAAGCAATATGGGAATACATGGGCCTGTTTGGGG 2520

Qy   2525 CATCAAACCACACAGTTTCCTATTGGGGCCTGTTAAGTGATGATGGGGAGAGAGGATGGAGA 2584
Db   2521 CATCAAACCACACAGTTTCCTATTGGGGCCTGTTAAGTGATGATGGGGAGAGAGGATGGAGA 2580

Qy   2585 ACAACTTTCACTGCGCGTGTCTCAGGGGAGGAGGAACCTGATGCACTGGGTCCCGTCCCTGC 2644
Db   2581 ACATCTTTCACTGCCCGGTGTCTCAGGGGAGGAGGGAACCTGATGCACTGGGTCCCGTCCCTGC 2640

Qy   2645 AGCTGCTGCTGCAGCCCACACTCTCTGGAGTCCTCCACTGCCTTGGTACCGAGACTCGGAACA 2704
Db   2641 AGCTGCTGCTGCAGCCCACACTCTCTGGAGTCCTCCACTGCCTTGGTACCGAGACTCGGAACA 2700

Qy   2705 AAAGGTTCCTGACACAGAGTGATGGGCCATTTCGGAAGAAATGGGGCATTGGGAGAAACAG 2764
Db   2701 AAACGTTCCTGACACAGAGTGATGGGCCATTTCGGAAGAAATGGGGCATTGTGTTAGAAACAG 2760

Qy   2765 ACATGGGAGCTCTTAGTGTTGCACTTTCTGCATTAAATTCAGCCGCGCCACGGTGAAGAGCTTC 2824
Db   2761 ACATGGGAGCTCTTAGTGTTGCACTTTCTGCATTAAATTCAGCCGCGCCACGGTGAAGAGCTTC 2820

Qy   2825 AGCTGAATTGAGGGGCAGCAGCCACAGAGATCAACATGGAGGCCCCACCATGGGTAGTCCTGTTCA 2884
Db   2821 AGCTGATTGAGGGGCAGCAGCCACAGAGATCAACATGGAGCCCCCACCATGGGTAGTCCTGTTCA 2880

Qy   2885 GGTGGGGTCCCAGTCACAGAGATGCCTATTGGCAGAATTCTCTTCTCTCGGTCCTCCTCAAAGGTCACCA 2944
Db   2881 GGTGGGGTCCCAGTCACAGAGATGCCTATTGGCAGAATTCTCTTCTCTCGGTCCTCCTCAAAGGTCACCA 2940

Qy   2945 GAAACCTGAAGAGACTGGGACCTAAGTGGAAAACTCGCTAAGTGAAGTC 3000
Db   2941 GAAACCTGAAGAGACTGGGACCTAAGTGGAAAACTCGCTAAGTGAAGTC 3000

Qy   3005 TTTGTAAGACCCTGAGACGCCCCTGCCTCCTTGGGTGGGGCTGGCTGTG 3064
Db   3001 TTTGTAAGACCCTGAGACGCCCCTGCCTCCTTGGGTGGGGCTGGCTGTG 3060

Qy   3065 GCCTCACAGCTGAGGACTGCAAGGACCTTGCCTTTGGGGTGAGAGCCAACCAGACCCTGA 3124
Db   3061 GCCTTCACAGCTGTGAGGACTGCAAGGACCTTGCCTTTGGGGCTGAGAGCCAACCAGACCCTGA 3120

Qy   3125 CCGAGCTGGACCTGACTGAGCTTCAAATGTGCTCACGGATGGCAAAAAACCTTTGCCAGA 3184
Db   3121 CCGAGCTGGACCTGACTGAGCTTCAAATGTGCTCACGGATGGCAAAAAACCTTTGCCAGA 3180

Qy   3185 GACTGAGACAGCCCAGGCTGCAAGCTACAGGGACTGCAGCTGGCTCGTCGTTGGCCTTCACGT 3244
Db   3181 GACTGAGACAGCCCAGGCTGCAAGCTACACGGGACTGCAGCTGGCTCGTCGTTGGCCTTCACGT 3240

Qy   3245 CTGACTGCTGCCAGGAACCTGGCCTCTCGTTGGCGTGCTGTTAGTCCAGCGACTGAAGAGGCTAG 3304
Db   3241 CTGACTGCTGCCAGGAACCTGGCCTCTCGTTGGCGTGCTGTTAGTCCAGCGACTGAAGAGGCTAG 3300

Qy   3305 ACCTGCAGCAGGAACAACTGGATGACGTTGGGCGTGGCCAACCCTGA 3364
Db   3301 ACCTTGCAGCAGGAACAACTGGATGACGTTGGGCGTGGCCAACCCTGA 3360

Qy   3365 ATCCTGCCCTGGCAAACTCATTACGCCCTGGGCTGGCTGTGGTGGATGATGA 3424
Db   3361 ATCCTGCCTGGCAAACTCATACGCCCTGGGCTGGCTGTGGTGGATGATGA 3420

Qy   3425 GGCAGGAACTGAGGGCCCTGGACAGGCAGAGAAACCTCAGCTCATTCAGTCAGAGGGA 3484
Db   3421 GGCAGGAACTGAGGGCCCTGGACAGGCAGAGAAACCTCAGCTCATTCAGTCAGAGGGA 3480

Qy   3485 AACCAAAGTGTGATGACCCTACTGAGGGGCCTGGATACGGGGATGAGTAATAGCACAT 3544
Db   3481 AACCAAAGTGTGATGACCCCTACTCAGAGGGGCCTGGATACGGGGATGAGTAATAGCACAT 3540

Db   3545 CCTCACTCAAGCGCCAGAGACTCGGATCAGAGAGAGGGCGGCTTCCCATGTTGCTCAGGCTA 3604
```

```
Db  3541 CCTTCACTCAAAGCCGGCAGAGACTCGGACTCAGAGAGAGGGCGGCCTTCCCATGGTTGCTCAGGCTA
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3605 ATCTCAAACTCCTTGGACGTGAGCAAGATCTTCCCAATTTGCTTGACATTGACAGAGGAAAGCT 3664
Db  3601 ATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATTTGCTTGACATTGACAGAGGAAAGCT 3660
Qy  3665 CCCCAGAGGTAGTTACCGGTGGAACTCTTGTTGCGTTGCGTGCCTTCTCCTTCTGTCAAGGGACC 3724
Db  3661 CCCCAGAGGTAGTTACCGGTGGAACTCTTGTTGCGTTGCGTGCCTTCTCCTTCTGTCAAGGGACC 3720
Qy  3725 TGCATACCGAAAGCCTTTTGGGGACTGACGATGACTTCTGAGGACCCCACACAGGGGCCCTGGGCTA 3784
Db  3721 TGCATACCGAAAGCCTTTTGGGGACTGACGATGACTTCTGAGGACCCCACACAGGGGCCCTGGGCTA 3780
Qy  3785 CTGAGGTAGTTGACAAAGAAAAGAAGAACTTGTACCGAGTTCACTTCCCGTGTAGCTGGCTCCTT 3844
Db  3781 CTGAGGTAGTTGACAAAGAAAAGAAGAACTTGTACCGAGTTCACTTCCCGTGTAGCTGGCTCCTT 3840
Qy  3845 ACCGCTGGGCCCAACACGGGGTCTCTGCTTTGTGATGAGAGAGGGGTGACCGTTGAGATTG 3904
Db  3841 ACCGCTGGGCCCAACACGGGGTCTCTGCTTTGTGATGAGAGAGGGGTGACCGTTGAGATTG 3900
Qy  3905 AATTCTGTGTGTGTGGGGACCAGTTCCTGGGGTGAGATCAACCCCACACACTGATGGTGG 3964
Db  3901 AATTCTGTGTGTGTGGGGACCAGTTCCTGGGGTGAGATCAACCCCACACACTGATGGTGG 3960
Qy  3965 CACGGGGCCTTCTGCTCGGACATCAAGCCTGAGCCTGTGGAAGCTGTGGCCACTTCCCTC 4024
Db  3961 CACGGGGCCTTCTGCTCGGACATCAAGCCTGAGCCTGTGGAAGCTGTGGCCACTTCCCTC 4020
Qy  4025 ACTTTGTGGCTCTCCAAGGGGGGCCAGTGGACACATCCCGTTCCAAATGGCCCACTTTA 4084
Db  4021 ACTTTGTGGCTCTCCAAGGGGGGCCAGTGGACACATCCCGTTCCAAATGGCCCACTTTA 4080
Qy  4085 AAGAGGAGGGGGATGCTCTCTGGAGAAGCCAGCCAGGGTTCCATCACATACTTTCTG 4144
Db  4081 AAGAGGAGGGGGATGCTCTCTGGAGAAGCCAGCCAGGGTTCCATCACATACTTTCTGG 4140
Qy  4145 AAAAACCCCCAGCTTCTCCCCCTTGGGAGTCCTCTGAAAATGATCCATAATGCCCTGGCT 4204
Db  4141 AAAAACCCCCAGCTTCTCCCCCTTGGGAGTCCTCTGAAAATGATCCATAATGCCCTGGCT 4200
Qy  4205 TCATTCCCGTCACCCTCGTGTGGTTGCTTACCACGCGGTCCATCCTGAGGAAGTCACCT 4264
Db  4201 TCATTCCCGTCACCCTCGTGTGGTTGCTTACCACGCGGTCCATCCTGAGGAAGTCACCT 4260
Qy  4265 TCCACCCTCTACCTGATCCCAAGTGACTGACTGCTTCCAATCGAGAGGCTCTGCTATC 4324
Db  4261 TCCACCCTCTACCTGATCCCAAGTGACTGCTTCCAATCGGAGAGGACTGCTTCTGCTATC 4320
Qy  4325 GAAGCCCTGGAGGAAGCAGCCAGCTGTTCTCGGAGTTCTACGTTGGGCCACTTGGGATCAGGGA 4384
Db  4321 GAAGCCCTGGAGGAAGCAGCCAGCTGTTCTCGGAGTTCTACGTTGGGCCACTTGGGATCAGGGA 4380
Qy  4385 TCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTGGTTGGTGGGAGGGCCTTGGTTGGTGAAAC 4444
Db  4381 TCAGGCTGCAAGTGAAAGACAAGAAAGATGAGACTCTGGTTGGTGGGAGGGCCTTGGTTGGTGAAAC 4440
Qy  4445 CAGGAGAGATCTCATGCCTGCAAACTACTCTGATCCCTCCAGCCCGCATAGCCGTACCTTCAC 4504
Db  4441 CAGGAGAGATCTCATGCCTGCAAACTACTCTGATCCCTCCAGCCCGCATAGCCGTACCTTCAC 4500
Qy  4505 CTCTGGATCTGCCCCCGCCAGTTGTGCTGCGCACTTTGTTGTGAGGACCAGCCTGA TAGCCC 4564
Db  4501 CTCTGGATCTGCCCCCGCCAGTTGTGCTGCGCACTTTGTTGTGAGGACCAGCCTGA TAGCCC 4560
Qy  4565 GAGTGACATCGGTGGAGGTTGTCTTGGACAAACTGCATGGACCAGGTGCTGAGCCAGGAGC 4624
Db  4561 GAGTGACATCGGTGGAGGTTGTCTTGGACAAACTGCATGGACCAGGTGCTGAGCCAGGAGC 4620
Qy  4625 AGTACGAGAGGGGTGCTGGCTGAGAACACGAGGCCCAGCAGATGCGGAAGCTGTTCAGCT 4684
Db  4625 AGTACGAGAGGGGTGCTGGCTGAGAACACGAGGCCCAGCAGATGCGGAAGCTGTTCAGCT 4684

Db  4621 AGTACGAGAGGGGTGCTGGCTGAGAACACGAGGCCCAGCAGATGCGGAAGCTGTTCAGCT 4680
Qy  4685 TGAGCCAGTCCTGGGACCGGAAAGTGCAAAAGATGGCAACTCTACCAAGCCCTGAAGGAGACC 4744
Db  4681 TGAGCCAGTCCTGGGACCGGAAAGTGCAAAAGATGGCAACTCTACCAAGCCCTGAAGGAGACCC 4740
Qy  4745 ATCCTCACCTCACTTGTCATTTATGAAGTGGGAGGGCAGCAAAAAAGGGGACTCCTGCCCACTCA 4804
Db  4741 ATCCTCACCTCATTATGAAGTGGGAGGGCAGCAAAAAAGGGGACTCCTGCCCACTCA 4800
Qy  4805 GCAGCTGAAGTATCAACACCAGCACGCCCTGACCCTTGAGTCCTGGCTTTGGCTTGTGACCCTTCT 4864
Db  4801 GCAGCTGAAGTATCAACACCAGCACGCCCTTGAGTCCTGGCTTTGGCTTGTGACCCTTCT 4860
Qy  4865 TTGGGTCTCAGTTTCTTCTCTCGCAAACAACAAGTTGCCATCTCGGTTTGCCTTCCTCCAGCACTAA 4924
Db  4861 TTGGGTCTCAGTTTCTTCTCTCGCAAACAACAAGTTGCCATCTCGGTTTGCCTTCCTCCAGCACTAA 4920
Qy  4925 AGTAATGGAACTTTGATGATGCCTTTGCTGGGAGTAATGTTGGTCCAATGCCAGGGATGCCAC 4984
Db  4921 AGTAATGGAACTTTGATGATGCCTTTGCTGGGAGTAATGTTGGTCCAATGCCAGGGATGCCAC 4980
Qy  4985 AGGGGGCCCCCAGTCCAGGTGGCCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5044
Db  4981 AGGGGGCCCCCAGTCCAGGTGGCCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5040
Qy  5045 AGACCCCTGCAGACCCTCATAGAGCCTTCATCTGGTGGCGCCACAGACCTAGAGCCC 5104
Db  5041 AGACCCCTGCAGACCCTCATAGAGCCTTCATCTGGTGGCCACAGACCTAGAGCCC 5100
Qy  5105 TCCGGATCCCATCCAGGGGCCAAAGAGGGGAATAGGAGGAGGGACCATTTGCCTCTGGC 5164
Db  5101 TCCGGATCCCATCCAGGGGCCAAAGAGGAAATAGGAGGAGGGACCATTTGCCTCTGGGC 5160
Qy  5165 TGTGTCACAGGGTGAGCCCAAAAATTGGGGTTCAGGGTGGAGGGCCACGTGGATTCTTGG 5224
Db  5161 TGTGTCACAGGGTGAGCCCAAAAATTGGGGTTCAGGGTGGAGGGCCACGTGGATTCTTGG 5220
Qy  5225 CTTTGTTACAGGAGAAGATCTACAAGAGCCAACACAGAGTAAAAGTGGGAAGGAAGTTTATTC 5284
Db  5221 CTTTGTTACAGGAGAAGATCTACAAGAGCCAACACAGAGTAAAAGTGGGAAGGAAGTTTATTC 5280
Qy  5285 AGAAAAATAAAGGAGTAATCACAGGCTCTCTTTAGAATTTGTCTAGCAGCTTTCCAGTTTTTTA 5344
Db  5281 AGAAAAATAAAGGAGTAATCACAGGCTCTCTTTAGAATTTGTCTAGCAGCTTTCCAGTTTTTA 5340
Qy  5345 CCAGAAAAACCCCTATAAAATTAAAATATTTTTTACTTAAAATTAAGAATTAAAAATACAA 5404
Db  5341 CCAGAAAAACCCCTATAAAATTAAAATATTTTTTACTTAAAATTAAGAATTAAAAA 5400
Qy  5405 AAAA 5408
Db  5401 AAAA 5404
```

RESULT 3

```
LOCUS       AX459865        5100 bp    DNA     linear   PAT 08-JUL-2002
DEFINITION  Sequence 8 from Patent WO0240668.
ACCESSION   AX459865
VERSION     AX459865.1  GI:21725641
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tschopp,J. and Martinon,F.
  TITLE     Proteins and dna sequences underlying these proteins used for
            treating inflammations
  JOURNAL   Patent: WO 0240668-A 8 23-MAY-2002;
            Apotech Research and Development Ltd. (CH)
FEATURES             Location/Qualifiers
```

source   1. .5100
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /note="NALP1.cdna"

BASE COUNT   1251 a   1374 c   1406 g   1069 t
ORIGIN

Query Match   89.7%; Score 4882.4; DB 6; Length 5100;
Best Local Similarity   99.5%; Pred. No. 0; Mismatches 26; Indels 0; Gaps 0;
Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

RESULT 4

```
LOCUS       AF298548                5100 bp    mRNA    linear   PRI 19-FEB-2002
DEFINITION  Homo sapiens caspase recruitment domain protein 7 mRNA, complete
            cds.
ACCESSION   AF298548
VERSION     AF298548.1  GI:10198208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5100)
  AUTHORS   Bertin,J. and DiStefano,P.S.
  TITLE     The PYRIN domain: a novel motif found in apoptosis and inflammation
            proteins
  JOURNAL   Cell Death Differ. 7 (12), 1273-1274 (2000)
  MEDLINE   21169419
   PUBMED   11270363
REFERENCE   2  (bases 1 to 5100)
  AUTHORS   Bertin,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Neurobiology, Millennium Pharmaceuticals
            Inc., 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1..5100
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             160..4449
                     /note="CARD7; nucleotide-binding site/leucine-rich repeat
                     family member (NBS/LRR); contains PYRIN and CARD domains"
                     /codon_start=1
                     /product="caspase recruitment domain protein 7"
                     /protein_id="AAG15254.1"
                     /db_xref="GI:10198209"
                     /translation="MAGGAWGRLACYLEFLKEEELKEFQLLLANKAHSRSSSGEETPAQ
                     PEKTSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQAEGAQHSPSFPYSPSE
                     PHLGSPSQOPTSTAVLMPWIHELPAGCTCGSERRVLRQLPDTSGRRWREISASHLYQAL
                     PSSPDHHESPSQESSPNAPFSTAVLGSWSGSSPPQSSLAPRREQEAPGTQWPLDETSGIYYTE
                     IRERERKSEKGRPPWAAVGTFPPQAHSSLDPHHHPWEPSVRESLCLDTSTWPWKNEDFNQ
                     KFTQLLLQRPHPRSCPDLVKRSWPDVVEBNRGHLIEIRDLFGPGLDQTTLLP
                     AAGIGKSTLARQVKEANGRGQLYEGDRRQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPRRLLFILDGVDEPGWVLQEPSSELCLHWSQOPPADALLGSLLGKTILP
                     EASFLIARTRIALQNIIPSLEQARWEVLGFSESSSRKEYFRYFTDDRQAIRAFRLVK
                     SNKELWALCLVPWNSWLACTCLMQOMRKEKLTLTSKTTTLCLHYLAQALQOPLGP
                     QLRDLCSLAAAGIWQKKTLFSPDDLRKHGLDGAIISTFLRMGILQEPIPLSYFIHL
                     CFQEFPAAMSVYLEDEKGRGKHSNCIIDLEKTLEAVGIHGLGQAASTRFLLGLLSDEG
                     EREMENIFHCRLSQGRNIMOWPSLQLLLQPHSLESHLCLYETRNKTFLTQVMAHFEE
                     MGMCVEIDMELLVCTFCIKFSRHVRKLQLIEGRQHRSPSMVVLFRWVDAYWQ
                     ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPCLLETIRLAGCGLTAEDCKD
                     LAFGLRANQTLTELDLSFNVLMDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDL
                     ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDENRQELR
                     ALEQEKPQLLIFSRRKPSVMTPIEGLDTGEMSNSTSSLKRQRLGSSERAASHVAQANLK
                     LLDVSKIFPIABIAEESSPEVVPEVLCVFSPASQGDLHTKPLGTDDDFWGPFGPVAT
                     EVVDKEKNLYRVHFPVAGSYRWPNTGLICFVVREAAVTVEIBFCVWDQFLGEINPQHSWM
                     IVLENFSSPLGVLLRMIHNALRFIPVTSVVLLYHRLHPEEVTFHLYIIPSDCSIRKE
                     VAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQVAHFEKEGMLLEKPARVELHH
                     LELCYRSPGEDQLFSEFYGHLGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP
                     ACIAVPSPLDAPQLLHVDQYREQLIARVTSVEVVLDXLHGQVLSQEQYERVLAENTR
                     PSQMRKLFSLSQSWDRKCKDGLIYQALKETHPHLIMELMWEKGSKKGLLPLSS"
```

```
BASE COUNT     1251 a     1374 c     1406 g     1069 t
ORIGIN
```

```
Query Match     89.7%;  Score 4882.4;  DB 9;  Length 5100;
Best Local Similarity   99.5%;
Matches 4898;  Conservative    0;  Mismatches   26;  Indels   0;  Gaps   0;
```

```
Qy         516 GACAGAGAATCGGCTCGGCGAGCCTCGGGGCGCGCCTTGCCTCGTTACTTGGAGTTCCTTC 575
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         153 GAACAAGAGGGGCTGCGCGAGCCTCGGGGCGCCCTGTTACTTGGAGTTCCTGAAGAA 212

Qy         576 GGAGGGAGCTGAAGGAGTTCCAGCTTCGTCGCCCAATAAAGCGCACTCCAGGAGAGCTCTTC 635
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         213 GGAGGAGCTGAAGGAGTTCCAGCTTCTGCGCCACTCCAGGAGGTCTTC 272

Qy         636 GGGGTGAGACACCGCTCAGCCAGAGGAAGACGAGTGGCATGGAGGTGGCCTTCGACCTGGT 695
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         273 GGGGTGAGACACCGCTCAGCCAGAGAAGACGAGTGGCCTCGTTACCTGGT 332

Qy         696 GGCTCAGTATGGGGAGCAGCGGGCCTGGGACCTTAGCCCTTCCATACCTGGGAGTCAGATGGG 755
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         333 GGCTCAGTATGGGGAGCAGCGGGCCTAGCCCTTCCATACCTGGGAGTCAGATGGG 392

Qy         756 GCTGAGGTCACTGTGCGCCCAGACCCAGACCCCAGGAAGGGCAGGCCCACTCTCCCTCATTCCCCTA 815
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         393 GCTGAGGTCACTGTGCGCCCAAGCCCCAGGAAGGGCAGGCCCACTCTCCCTCATTCCCCTA 452

Qy         816 CAGCCCAAGTGAAACCCCACCTGGGGT 875
               |||||||||||||||||||||||||||
Db         453 CAGCCCAAGTGAACCCCACCTGGGGTCTCCCAGACGAGAAGGGCAGCAGTGCTAAT 512

Qy         876 GCCCTGGATCCATGAATTGCCGCGGGGTGCACCCAGGGGCTCAGAGAGGAAGGGTTTGAG 935
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         512 ...
```

```
Qy        4716 TGGACTCTACCAAGCCCTGAAAGAGAGACCCATCCTCACTTCATTATGTGGAACTCTGGGAGAA 4775
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4353 TGGACTCTACCAAGCCCATCCTCACTTCATTATGTGGAACTCTGGGAGAA 4412

Qy        4776 GGGCCAGCAAAAAGGAGACTTCCTGCCACTCAGCAGCTGAAGCTACAACACCAGCCCTTGACC 4835
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4413 GGGCCAGCAAAAAGGAGACTTCCTGCCACTCAGCTACAACACCAGCCCTTGACC 4472

Qy        4836 CTTGAGTCCTGGGCTTTGGCTGACCCCTTCTTTCTTGCAAACAAG 4895
               |||||||||||||||||||||||||||||||||||||||||||||
Db        4473 CTTGAGTCCTGGGCTTTGGCTGACCCCTCAGTTTCTTCTGCAAACAAG 4532

Qy        4896 TTGCCATCTGGTTTTGCCTTCCAGCACTAAAAGTAATGGAACTTTGATGATGCCTTTGCTGG 4955
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4533 TTGCCATCTGGTTTTGCCTTCCAGCACTAAAAGTAATGGAACTTTGATGATGCCTTTGCTGG 4592

Qy        4956 GCATTAGTGTCCATGCCAGGGGATGCCACAGGGGGGGCCCCAGTCCAGGTGGCCTAACAGCA 5015
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4593 GCATTAGTGTCCATGCCAGGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACAGCA 4652

Qy        5016 TCTCAGGGAAATGCCATCTGGAGCTGGCAAGAGACCCCTCGCAGAGACCTCATAGAGCCTCACT 5075
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4653 TCTCAGGGAAGTGCCATCTGGAGCTGGCAAGACCCTCAGAGACCTCATAGAGCCTCATCT 4712

Qy        5076 GGTGGCCCACAGCAGCCCAAGCCCTAGAGCCCTTCCGGATCCCATCCAGGGGCGCAAAAGAGGAATA 5135
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4713 GGTGGCCCACAGCAGCCCAAGCCTACAGAGCCCTTCCGGATCCCATCCAGGGGCGCAAAAGAGGAATA 4772

Qy        5136 GGAGGGACATGGAACCATTGCCTCTGGCGTCGTGTGTGTCACGAGAGGGTGAGCCCCAAAATTGGGGT 5195
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4773 GGAGGGACATGGAACCATTGCCTCTGGCGTGTGTCACGAGGGTGAGCCCCAAAATTGGGGT 4832

Qy        5196 TCAGCGTGGGAGGGCCGCCGGATCCTTCTTGGCTTTGTACAGGAGGAAGATCTTCAGATGAGCAGCC 5255
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4833 TCAGCGTGGGAGGGCCGGATCCTTCTTGGCTTTGTACAGGAGAAGATCTACAAAGAGAGCAGCC 4892

Qy        5256 AACAGAGTAAAGTGGGAAGGAAGTTTATTCAGAAAATAAAGGAGAGTAATCACAGCTCTTTTAG 5315
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4893 AACAGAGTAAAGTGGGAAGGAAGTTTATTCAGAAAATAAAGGAGTATCACTGCTCTTTTAG 4952

Qy        5316 AATTTGTCTAGCAGGCTTTCCCAGTTTTTACAGAAAACCCTATAAATTAAATTAAAATTTTT 5375
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4953 AATTTGTCTAGCAGACTTCCAGTTTTTACAGAAAAACCCTATAAATTAAAATTTTT 5012

Qy        5376 ACTTAAATTTAAGAATTAAAAAATACAAAAAAGAGAAATTGAAAATAAGACGGAATAAGAA 5435
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5013 ACTTAAATTTAAGAATTAAAAAATACAAAAAAAATGAAAATAAGGAATAAGAA 5072

Qy        5436 GTTTA 5439
               |||||
Db        5073 GTTTA 5076
```

RESULT 5

LOCUS   AX207661   5122 bp   DNA   linear   PAT 31-AUG-2001
DEFINITION Sequence 29 from Patent WO0157085.
ACCESSION   AX207661
VERSION   AX207661.1 GI:15422343
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baughn,M.R., Au-Young,J. and Yue,H.
  TITLE   G-protein coupled receptors
  JOURNAL   Patent: WO 0157085-A 29 09-AUG-2001;
            Incyte Genomics, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..5122
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 4250893CB1"
BASE COUNT   1184 a   1460 c   1401 g   1077 t
ORIGIN

Query Match   87.5%; Score 4765.6; DB 6; Length 5122;
Best Local Similarity   96.9%; Pred. No. 0; Mismatches 24; Indels 133; Gaps 2;
Matches 4933; Conservative   0; Mismatches   24; Indels 133; Gaps 2;

Qy   35 GAGCCAGCAGCCGGGGGCCTCCACTCTGGGGTTCTGAAAGCCATTCCCTGCCTGCGCGCTC 94
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GAGCCAGCAGCCGGGGGCCTCCACTCTGGGGTTCTGAAAGCCATTCCCTGCCTGCGCGCTC 60

Qy   95 CTCCCACCCACTCTTCTCAGCGTTCGAGCTCCAGGGTTGATCTCAGGAGTCCAGGACC 154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCCCACCCACTCTTCTCAGCGTTCGAGCTCCAGGGTTGATCTCAGGAGTCCAGGACC 120

Qy   155 CAGGAGAGGGAGAGAAATCTGAGGAACACAGACAGTGAGGCGTTGCCCACCCCATCTCCC 214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGGAGAGGGAGGAAGAAATCTGAGGAACACAGACAGTGAGGCGTTGCCCACCCCATCTCCC 180

Qy   215 GTCACCACAACATCTCCCCTCACCCTACCCTCCCTGCCCTGGCCCTTGGACCCCATCCCAGGAC 274
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTCACCACAACATCTCCCCTGTAACACCCTCCCTGCCCTGGCCCTTGGACCCCATCCCAGGAC 240

Qy   275 CTCCCTATCAGCTGACTTCTTCCAGTGTCTTGCAGGGCCCCTGGCTCCCTCCCTCCCCCT 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTCCCTATCAGCTGACTTCTTCCAGTGTCTTGCAGGGCCCCTGGCTCCCTCCCTCCCCCT 300

Qy   335 GGCTTTTCCTACCACTCCCCCCTCTATCGGCGGTCTATCTGTAGGTGCCCTGGGGATTTTATAA 394
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGCTTTTCCTACCACTCCCCCCTCTATCGGCGGTCTATCTGTAGGTGCCCTGGGGATTTTATA 360

Qy   395 AACTGGGTTCCGAAATGCTGAATTAAGGAGACGGTTAAGAGCCCAAGGCAAAGGACAGCACACTGTT 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AACTGGGTTCCGAAATGCTGAATTAAGGAGACGGTTAAGAGCCCAAGGCAAAGGACAGCACACTGTT 420

Qy   455 CTCTGCCTGCCTTGGATACCCTCACCACCTGGGAACATCCCCCAGACACCACCCTCTTTAACTCCG 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTCTGCCTGCCTTGGATACCCTCACCACCTGGGAACATCCCCCAGACACCACCCTCTTTAACTCCG 480

Qy   515 GGACAGAGATGGCTGGGCGGAGACCTGGGGCCGGGCTGTTACTTGGGATTCCTGAAGA 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GGACAGAGATGGCTGGGCGGGGGAGACCTGGGGCCGCCTGGCCTGTTACTTGGGATTCCTGAGA 540

Qy   575 AGGAGGAGGCTGAAGGAGGCTGAAGTTCAGCTTCTGCTCCGCCAATAAAGCGCACTCCAGGAGCTCTT 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGGAGGAGGCTGAAGGAGGTTCAGCTTCTGCTCCGCCAATAAAGCGCACTCCAGGAGCTCTT 600

Qy   635 CGGGGTGAGACACCCCGCTCAGCCTCAGAGCAGAGAAGACCAGTGGCATGGAGGTGGCCTTGTTCCTGG 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CGGGGTGAGACACCCCGCTCAGCCTCAGAGCAGAGAAGACCAGTGGCATGGAGGTGGCCTTGTTCCTGG 660

Qy   695 TGGCTCAGTATGGGGAGCAGCGGCGGGCCCTGGGAGACCTAGCCCTTCCATACCTGGAGGACGATGG 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TGGCTCAGTATGGGGAGCAGCGGCGGGCCCTGGGAGACCTAGCCCTTCCATACCTGGGGACGATGG 720

Qy   755 GGCTGAGGTCACTGTGCGCCCAAGCCCAGAGGAAGGGGACCAGGCCCACTCTCCCCTCATTCCCT 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GGCTGAGGTCACTGTGCGCCCCAAGCCCCAGAGGAAGGGGAGCCACTCTCTCCCTCATTCCCT 780

Qy   815 ACAGCCCAAGTGAACCCCACCTGGGGTCTCCCAGCCCAACCCACCTCCACCTCCACCGCAGTGCTAA 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ACAGCCCAAGTGAACCCCACCTGGGGTCTCCCAGCCCAACCCACCTCCACCGCAGTGCTAA 840

Qy   875 TGCCCTGGATCCATGAATTGCCGGGCGGGGGTGCACCCAGGCTCAGAGGAAGGAGGTTTTGA 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCCTGGATCCATGAATTGCGGGCGGGGGTGCACCCAGGCTCAGAGGAAGGAGGTTTTGA 900

Qy   935 GACAGCTGCCCTGACACATCTGACGCGCGCTGAGAGAAATCCTCTGCCTCACTCCTCTACC 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GACAGCTGCCCTGACACATCTGACGCGCCCTGAGAGAGAAATCTCTGCCTCACAGTCTACC 960

Qy   995 AAGCTCTTCCAAGCTCCCCAGACCCATGAGTCTCCAAGCCCAGAGTCACCCAAGCCCCCA 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AAGCTCTTCCAAGCTCCCCAGACCCATGAGTCTCCAAGCCCAGAGTCACCCAAGGCCCCA 1020

Qy   1055 CATCCACAGCAGTGCTCGGGGAGCTGGGGATCCCACCTCAGCCTCAGCCCTAGCACCCAGAG 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCCACAGCAGTGCTGGGGGATCCCACCTCAGCCTCAGCCTAGCACCCAGAG 1080

Qy   1115 AGCAGGAGGCCTCTGGGGACCCAAATGGGCCTCTGGATGAATGAAACGTCAGGAATTTACTACACAG 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGCAGGAGGCTCTGGGGACCCAAATGGGCCTCTGGATGAAACGTCAGGAATTTACTACACAG 1140

Qy   1175 AAATCAGAGAAAAGGAGAGAGAGAGAGAGAGAAATCAGAGAAAAGGGAGGGCGGGTGG 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AAATCAGAGAAAAGAGAGAGAGAGAAATCAGAGAAAAGGCAGGGCGGGTGG 1200

Qy   1235 TAGGAACGGCCCCCACAGGCGCACACCAGCCTACAGCCCCCACCACCCATGGGGGAGCCTT 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TAGGAACGCCCCCACAGGCGCACCAGCCTACAGCCCCCACCACCCATGGGGGAGCCTT 1260

Qy   1295 CTGTGAGAGGAGAGCCCTCTGTTCCACATGGCCCTGGGAAAAAATGAGGATTTTAACCAAAAAT 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTGTGAGAGAGAGCCCTCTGTTCCACATGGCCCTGGGAAAAAATGAGGATTTTAACCAAAAAT 1320

Qy   1355 TCACACAAGCTGCTTACTTCTACAAAGACTCACCCGAGAGACTCCCTGGTTCAAGA 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCACACAAGCTGCTTACTTCTACAAAGACTCACCCGAGAGACTCCCTGGTTCAAGA 1380

Qy   1415 GAAGCTGGCCTGATTATGTGGAGGAATGGGAGGACATTTAATTGAGATCAGAGACTTTAT 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GAAGCTGGCCCTGATTATGTGGAGGAATCGGAGGACATTTAATTGAGATCAGAGACTTTAT 1440

Qy   1475 TTGGTCCCCAGGGCCTGGGATACCCAAGAAACCTCGCATAGTCATACTGCAGGGGTGCTGGAA 1534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTGGTCCCCAGGGCCTGGGATACCCAAGAAACCTCGCATAGTCATACTGCAGGGGTGCTGGAA 1500

Qy   1535 TTGGGAAGTCAACACTGCCAGGCAGGTGAAGAGGAAGCCTGGGGGAGGAGGCCAGCTGTATG 1594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ... 1594

```
Qy   3755 ACTTCTGGGGGCCCCCACGGGGGGCCTGTGGGGCTACTGAGGTAGTTGACAAAGAAAAGAAACTTGT 3814
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACTTCTGGGGGCCCCCACGGGGGCCTGTGGGGCTACTGAGGTAGTTGACAAAGAAAAGAAACTTGT  3780

Qy   3815 ACCGAGTTCACTTCCCTGTAGCTGGCTCCTACCGGTCTGCGGTCTCTCTGCTTTG            3874
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 ACCGAGTTCACTTCCCTGTAGCTGGCTCCTACCGGTCTGCGGTCTCTCTGCTTTG            3874

Qy   3875 TGATGAGAGAAGGCGGTGACCGTTGAGATTGAATTCCTGTGTGTGTGGGACCAGTTCCTGGGGTG  3934
          ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   3841 TGATGAGAGAAGGCGGTGACCGTTGAGATTGAATTCCTGTGTGTGTGGGACCAGTTCCTGGGTG   3900

Qy   3935 AGATCAACCCCACCAGCACAGCAGCTGGATGGCAGGGCCTTCGTGTTGCAGACATCAAGGCTGAGC 3994
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 AGATCAACCCCACCAGCACAGCAGCTGGATGGCAGGGCCTTCGTGTTGCAGACATCAAGGCTGAGC 3960

Qy   3995 CTGGAGCTGTGGAAGCTGTGTGCACCTCCACTTTGTGTGCTCTCCAAGGGGGGCCATGTGG      4054
          |||||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||
Db   3961 CTGGAGCTGTGGAAGCTGTGCACCTCCCACTTCCCAAGGGGGGCCATGTGG               4020

Qy   4055 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGGGGGATAGTCTCCCAAGAGAAGCCAG 4114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGGGGGATGCTCTTCCTGGAGAAGCCAG 4080

Qy   4115 CCAGGGTGGAGCTGCATCACATAGTTCTGGAAAAACCCAGCTTCTCCCCCTTGGGAGTCC        4174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 CCAGGGTGGAGCTGCATCACATAGTTCTGGAAAAACCCAGCTTCTCCCCCTTGGGAGTCC        4140

Qy   4175 TCCTGAAAAATGATCATAATGCCCTGCGCTTCATTCCCGTGCACCTCTGTGGTGTTGCTTT       4234
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TCCTGAAAAATGATCATAATGCCCTGCGCTTCATTCCCGTGCACCTCTGTGGTGTTGCTTT       4200

Qy   4235 ACCACCGCGTCCATCCTGAGGAAGTCACCTTCCACCTCTACCTGATCCCAAGTGACTGACTGCT   4294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 ACCACCGCGTCCATCCTGAGGAAGTCACCTTCCACCTCTACCTGATCCCAAGTGACTGACTGCT   4260

Qy   4295 CCATTCGG--------------------------------------------------------- 4302
          ||||||||
Db   4261 CCATTCGGAAGGCCATAGATGATCTAGAAATGAAAATTGCGAATTCCACGTTGTGCGAAATCCACAGC 4320

Qy   4303 ----------------------------------------------------------------- 4302
Db   4321 CACCCCCGCTGACCCCACTTTATATGGGCTGTCGTTACACTGTGTCTGGGGTCTGGTTCAG       4380

Qy   4303 ------------------------------------------------------------       4302
          
Db   4303 -----------------------------------------------------------       4302

Qy   4303 ------------AAGGAACTGGAGCTCTGCTATCGGAAGGCCCTGGAGAAGACC             4342
                      ||||||||||||||||||||||||||||||||||||||||||
Db   4381 GGATGCTGGAAAATACTCCCCAAGGAACTGGAGCTCTGCTATCGGAAGGCCCTGGAGAAGACC     4440

Qy   4343 AGCTGTGTTCTCGGAGTTCTACGTGGCCGATCAGGGGATCAGGGACTGTCAGGTCGCAAGTGAAAG 4402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 AGCTGTGTTCTCGGAGTTCTACGTGGCCGATCAGGGGATCAGGGACTGTCAGGTCGCAAGTGAAAG 4500

Qy   4403 ACAAGAAAAGATGAGAGACTCTCGGTGTGGGGGAGCCTTTGGTGGAAACCAGGAGATCTCATGCCTG 4462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 ACAAGAAAGATGAGAGACTCTCGGTGTGGGGAGGCCTTGGTGAAAACCAGGAGATCTCATGCCTG    4560

Qy   4463 CAACTACTCTCGATCCCTCCAGCCCCATAGCCGTACCTTCACCCTCTGGATGCCCCGCAGT       4522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 CAACTACTCTTGTTGACCCTCCAGCCCGACTTCACCCTCGGATGCCCCGCAGT              4620

Qy   4523 TGCTGCACTTGTGGACCAGTATCGATAGCCCGAGTGACGCTGAGCATCGGTCGGTGGAGG       4582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 TGCTGCACTTTGTGACCAGTATCGAGAGCCGAGTGACACTCGGTCGGTGGAGG              4680

Qy   4583 TTGTCTTGACAAAACTGCATGGACAGGTGCTGAGCGCCAGGAGCAGTACGAGAGGGTGCTGG     4642
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TTGTCTTGGACAAACTGCATGGAGACGTGCTGAGCGCCAGTACGAGAGGGTGCTGG          4740

Qy   4643 CTGAGAACACGAGGGCCCAGCCAGGAGCTGTTCAGCTTGAGCCAGTCCTGGGACC           4702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CTGAGAACACGAGGGCCCAGCCAGATGCGGAAGCTGTTCAGCTTGAGCCAGTCCTGGGACC     4800

Qy   4703 GGAAGTGCAAAGATGAGACTCTACCCAAGCCCCTGAAGGAGAGAGACCCATCCTCACCTCATTATGG 4762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GGAAGTGCAAAGATGAGACTCTACCCAAGCCCCTGAAGGAGAGAGACCCATCCTCACCTCATTATGG 4860

Qy   4763 AACTCTGGGAGGAGAGGGCAGCAAAAAGGGAGACTCCTGCCACTCAGCAGTGAGTATCAACA      4822
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 AACTCTGGGAGGAGAGGGCAGCAAAAAGGGAGACTCCTGCCACTCAGCAGTGAGTATCAACA      4920

Qy   4823 CCAGCCCTTGACCCTTGAGTCCTGGCTTTGCTGACACCTTCTTTGGGTCTCAGTTTCTTT       4882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 CCAGCCCTTGACCCTGAGTCCTGGCTTTGCTGACACCTTCTTTGGGTCTCAGTTTCTTT        4980

Qy   4883 CTCTGCAAACAACAAGTTGCCATCTGGTTTTGCCTTCAGCACTAAAGTAATGGGAACTTTGATG    4942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CTCTGCAAACAAGTTGCCATCTGGTTTTGCCTTCAGCACTAAAGTAATGGGAACTTTGAT       5040

Qy   4943 ATGCCTTTGCTGGGCATTATGTGTCCATGTGTCCA-GCCAGGGATGCACAGGGGGGGCCC      4992
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   5041 GATGCCTTGCTGGGCATTATGTGTCCA-GCCAGGGATGCACAGGGGGGGCCC              5089
```

ORIGIN

```
Query Match     78.8%;  Score 4290;  DB 9;  Length 4290;
Best Local Similarity  100.0%;
Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy    523 ATGGCTGGGGGAGCCTGGGGCCCGCCCTGGGCCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGGGGAGCCTGGGGCCCGCCCTGGGCCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 60
```

Qy CFQEFFAAMSYVLEDERGRGKHSNCIIDLEKTLEAYGIHGLFGASTTRFLLGLLSDEG
Db EREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLHCLYFTRNKTFLTQVMAHFEE
    MGMCVEFDMELLLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAVWQ
    ILFSVLKVTRNLKEIDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKQ
    IAFGLRANQTLTELDLSFNVLTDAGAKHLCEGLRHPACKLIRLGLDQTTLSEMRQELR
    ASVLSASPSLKELDLQQNLLDDVGVRLLCEGLKHPACKLIRLGLDQTTLSEMRQELR
    ALEQEKPQLLIFSRRKPSVMPTEGLDTGEMSNSTLSSLKRQRLIGSERAASHVAQANLK
    LLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGPTGPVAT
    EVVDKRENLYRVHFPVAGSYRWPNTGLCFVMREAVVTEIEFCPWDQFGEINPQHSHM
    VAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEEGMLLEKPARVELHH
    IVLENPSFPLGGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIRKE
    LELCYRSPGEDQLFSEFYVGHLGSGIRLQVKDKKDETLVWEALVKPGDLMPAITTLPP
    ARIAVPSPLDAPQLLHHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEYERVLAENTR
    PSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS"

```
Qy     61 CTGAAGGAGTTCCAGCTTTCTGCTCCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGAAGGAGTTCCAGCTTTCTGCTCCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120

Qy    643 ACACCCGCTCAGCCAGAGAAGACGAGTGGCATGGAGGTGGCCCTGGTTACTCTGGGTCTCAG 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACACCCGCTCAGCCAGAGAAGACGAGTGGCATGGAGGTGGCCCTGGTTACTCTGGGTCTCAG 180

Qy    763 TCACTGTGTGCGCCCAAGCCCAGGAAGGGGCCAGGACCCACTCTCCCTCATTCCCTACAGCCCA 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATGGGGACGAGCGGGGGCCCTGGGACCTAGCCCTTCATTCATCATCATTGGGAGCACGATGGGG 240

Qy    703 TATGGGGACGAGCGGGGGCCCTGGGACCTAGCCCTTCATTCATCATCATTGGGAGCACGATGGGG 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTGTGTGCGCCCAAGCCCAGGAAGGGGCCAGGACCCACTCTCCCTCATTCCCTACAGCCCA 300

Qy    823 AGTGAAACCCCACCTGGGGGTCTCCCAGCCAACCCACCCACCTCCACCCGACCAGGCTGCTAATGCCCTGG 882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGTGAAACCCCACCTGGGGGTCTCCCAGCCAACCCACCCACCTCCACCCGACCAGGCTGCTAATGCCCTGG 360

Qy    883 ATCCATGAATTGCCGGCCGGGGGTGCACCCAGGGGTCTCAGAGCCTCAGAGAGGAGGGTTTTTGAGACAGCTG 942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCCATGAATTGCCGGCCGGGGGTGCACCCAGGGGTCTCAGAGCCTCAGAGAGGAGGGTTTTTGAGACAGCTG 420

Qy    943 CCTGACACATCTGGACGCCGGCCTGGAGAGAAATCTTCTGCCTCACTCCTTCTACCAAGCTCTT 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTGACACATCTGGACGCCGGCCTGGAGAGAAATCTTCTGCCTCACTCCTTCTACCAAGCTCTT 480

Qy   1003 CCAAGCTCCCCAGACCATGAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCACCACCAATTCCACA 1062
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCAAGCTCCCCAGACCATGAGTCTTCCAAGCCAGGAGTCACCCAACGCCCCACCACCAATTCCACA 540

Qy   1063 GCAGTGCTCTGGGGACTCTGGGGGATCCCCACCTCAGCCGCCAGCCGCTAGCACTTCCAGACAGCAGGAG 1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCAGTGCTCTGGGGACTCTGGGGGATCCCCACCTCAGCCGCCAGCCGCTAGCACTTCCAGACAGCAGGAG 600

Qy   1123 GCTTCCTGGGACCCAATGGGCCTCTTGGGATGAAACGTCAGAGAATTTACTACAGACAGAAATCAGA 1182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCCTGGGACCCAATGGGCCTCTTGGGATGAAACGTCAGAGAATTTACTACAGACAGAAATCAGA 660

Qy   1183 GAAAGAGAGAGAGAGAAATCAGAGAAAAAGGCAGGCCCCCCATGGGGACAGGCCTTCGGTGGTGTAGGAACG 1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAAAGAGAGAGAGAGAAATCAGAGAAAAAGGCAGGCCCCCCATGGGGACAGGCCTTCGGTGGTGTAGGAACG 720

Qy   1243 CCCCCACACAGGCGCACACCAGCCCTACAGCCCCACACCACCCATGGGAGCGCTTCTGTGGTAAGA 1302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCCCACAGGCGCACACCAGCCCTACAGCCCCACACCACCCATGGGAGCGCTTCTGTGAGA 780
```

```
Qy   1303 GAGAGCCCTGTTCCACATGGGCCGCCCTGGGAAAAAAATGAGGATTTTAACCAAAAAAATTCACACAG 1362
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAGCCCTGTTCCACATGGGCCGCCCTGGGAAAAAAATGAGGATTTTAACCAAAAAAATTCACACAG 840

Qy   1363 CTGCTACTTCTACAAAAGACCTCACCCCAGAAGAGCCAAGATCCCTGGTCAAGAGAAGAGCTGG 1422
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGCTACTTCTACAAAAGACCTCACCCCAGAAGAGCCAAGATCCCTGGTCAAGAGAAGAGCTGG 900

Qy   1423 CCTGATTATGTGGAGGAGAATCGAGGACAGCAATTTAATTGAGATCAGAGACTTAATTTGGCCCA 1482
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTGATTATGTGGAGGAGAATCGAGGACAGCAATTTAATTGAGATCAGAGACTTAATTTGGCCCA 960

Qy   1483 GGGCCTGGGATACCCAAGAAACCTCGCATAGTCATACTGCACGGGGGCTGCTGGAATTGGGAAG 1542
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGGCCTGGGATACCCAAGAAACCTCGCATAGTCATACTGCACGGGGGCTGCTGGAATTGGGAAG 1020

Qy   1543 TCAACACTGGCCAGGCAGGTGAAGGAAGCCTGGGGGGAGGGGCCAGCTGTATGGGGACCGC 1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCAACACTGGCCAGGCAGGTGAAGGAAGCCTGGGGGGAGGGGCCAGCTGTATGGGGACCGC 1080

Qy   1603 TTCCAGACAGTGTCTTCTTCAGTGTCAGTGCAGAGAGACAGAGACAGCTCAAAGGTGGTGAGTCTC 1662
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTCCAGACAGTGTCTTCTTCAGTGTCAGTGCAGAGAGACAGAGACAGCTCAAAGGTGGTGAGTCTC 1140

Qy   1663 GCTGAGCTCATCGGAAAAAGATGGGACAGGCCACTCCGGGCTCCCATTAGACAGATCCTGTCT 1722
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCTGAGCTCATCGGAAAAAGATGGGACAGGCCACTCCGGGCTCCCATTAGACAGATCCTGTCT 1200

Qy   1723 AGGGCAGAGAGGGCTGCTCTTCATCCTCGATGGTGTGAGGCCAGGATGGGGTCTTGCAG 1782
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGGGCAGAGAGGGCTGCTCTTCATCCTCGATGGTGTGAGGCCAGGATGGGGTCTTGCAG 1260

Qy   1783 GAGCGGAGTTGTGAGCTCTTGTCTCTGCACTGGACCCACGCGGGGAGCGGATGCACTGCTG 1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GAGCGGAGTTGTGAGCTCTTGTCTCTGCACTGGACCCACGCGGGGAGCGGATGCACTGCTG 1320

Qy   1843 GGCAGTTTGCTTGGGGAAAACTATACTTCCTTCCTGATCACGGGCTCGGACC 1902
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGCAGTTTGCTTGGGGAAAACTATACTTCCTTCCTGATCACGGGCTCGGACC 1380

Qy   1903 ACAGCTCTGCAGAACCTCATTCCTTCTTCTTTGGGAGCAGGCACGTTGGGTAGAGGTCCTGGGG 1962
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACAGCTCTGCAGAACCTCATTCCTTCTTCTTTGGGAGCAGGCACGTTGGGTAGAGGTCCTGGGG 1440

Qy   1963 TTCTCTGAGTCCACCAGGAAGGAGGGAATTATTTCTACAGATGTTCACAGATAAAAGGCAAGCA 2022
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTCTCTGAGTCCACCAGGAAGGAGGGAATTATTTCTACAGATGTTCACAGATAAAAGGCAAGCA 1500

Qy   2023 ATTAGGAGCCCTTTAGGTTTGGTCAAATCAAAACAAAAGAGACTCTGGGGCCCTGTGTCTTGTGCCC 2082
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATTAGGAGCCCTTTAGGTTTGGTCAAATCAAAACAAAAGAGACTCTGGGGCCCTGTGTCTTGTGCCC 1560

Qy   2083 TGGGTGTCCTGGGCTGGCCTGCACTTGGCCTGATGCAGCGAGGAGGAAGAGGAAAAACTC 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TGGGTGTCCTGGGCTGGCCTGCACTTGGCCTGATGCAGCGAGGAGGAAGAGGAAAAACTC 1620

Qy   2143 ACACTGACTTCCAAGACCACCACAACCCTCTGTTACATTACCTTGCCCAGGCCTCCAA 2202
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACACTGACTTCCAAGACCACCACAACCCTCTGTTACATTACCTTGCCCAGGCCTCCAA 1680

Qy   2203 GCTCAGCCAATTGGGACCCCAGCCTCAGAGACCTCTCTCTGGCTGAGGGCATCTGG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCTCAGCCAATTGGGACCCCAGCCTCAGAGACCTCTCTCTGGCTGAGGGCATCTGG 1740

Qy   2263 CAAAAAAAAGACCTTTTCAGTCCAGATGACCTCAGGAAGCCATGGGTTAGATGGATGGGCCATC 2322
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCCATGGGTTAGATGGGCCATC 1800

Qy   2323 ATCTCCACCTTTCTTGAAAGATGGGTTATTCTTCAAGAGACCACCCATCCCTCTGAGTTACAGC 2382
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCTCCACCTTTCTTGAAAGATGGGTTATTCTTCAAGAGACACCCCATCCCTCTGAGCTACAGC 1860

Qy   2383 TTCATTCACCTCTGTTTCCAAGAGAGTTCTTTGCAGCAAATGTCCTATGTCTTGGGAGGAATGAG 2442
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2383 CCCCCACAGGCGCACACCAGCCCTACAGCCCCACACCACCCATGGGAGCGCTTCTGTGAGA 2442
```

```
Qy   4603  GGACAGGTGCTGAGCGAGCAGTAGCGAGAGGTGCTGGCTGAGGAACACGAGCCCCAGC  4662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081  GGACAGGTGCTGAGCCAGCCAGAGCTACGAGAGGGTGCTGGCTGAGGAACACGAGCCCAGC  4140

Qy   4663  CAGATGCGGAAGTGTTCAGCTTGAGCACTGCAGTCTCGGGACCGGAAGTGCAAAGATGGACTC  4722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141  CAGATGCGGAAGTGTTCAGCTTGAGCACTGCAGTCTCGGGACCGGAAGTGCAAAGATGGACTC  4200

Qy   4723  TACCAAGCCCTGAAGGAGGACCCATCCTCTCATTATTGGAACTTCTTCAGGAGGAAGGGCAGC  4782
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201  TACCAAGCCCTGAAGGAGGACCCATCCTCACCTCATTATTGGAAACTTCTTCAGGAGGAAGGGCAGC  4260

Qy   4783  AAAAAGGGGACTCCTCGCCACTCAGCAGCTGA  4812
           |||||||||||||||||||||||||||||||
Db   4261  AAAAAGGGGACTCCTCGCCACTCAGCAGCTGA  4290
```

                                                                                        R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 107 Row: j Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14719827.

```
RESULT 7
LOCUS          BC051787               5306 bp    mRNA    linear   PRI 14-MAY-2003
DEFINITION     Homo sapiens mRNA similar to death effector filament-forming
               Ced-4-like apoptosis protein (cDNA clone MGC:57544 IMAGE:5756099),
               complete cds.
ACCESSION      BC051787
VERSION        BC051787.1  GI:30705000
KEYWORDS       MGC.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 5306)
  AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
               Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
               Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
               Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
               Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
               Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
               Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
               Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
               Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
               McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
               Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
               Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
               Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
               Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
               Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
               Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
               Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
               Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE        Generation and initial analysis of more than 15,000 full-length
               human and mouse cDNA sequences
  JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED      12477932
REFERENCE      2  (bases 1 to 5306)
  AUTHORS      Strausberg,R.
  TITLE        Direct Submission
  JOURNAL      Submitted (01-MAY-2003) National Institutes of Health, Mammalian
               Gene Collection (MGC), Cancer Genomics Office, National Cancer
               Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
               USA
   MEDLINE
REMARK         NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT        Contact: MGC help desk
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Life Technologies, Inc.
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Sequencing Group at the Stanford Human Genome
               Center, Stanford University School of Medicine, Stanford, CA 94305
               Web site: http://www-shgc.stanford.edu
               Contact: (Dickson, Mark) mcd@paxil.stanford.edu
               Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
```

```
FEATURES             Location/Qualifiers
     source          1..5306
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:57544 IMAGE:5756099"
                     /tissue_type="Blood, adult leukocytes"
                     /clone_lib="NIH_MGC_118"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             553..4680
                     /codon_start=1
                     /product="Similar to death effector filament-forming
                     Ced-4-like apoptosis protein"
                     /protein_id="AAH51787.1"
                     /db_xref="GI:30705001"
                     /translation="MAGGAWGRLACYLFEFLKKEELKEFQLLLANKAHSRSSGGETPAQ
                     PEKTSGMEVASYLVAQVGEQRAWDLALHTWEQMGLRSLCAQAEGAGHSPSFPEYSPSE
                     PHLGSPSQGPTSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQAL
                     PSSPDHBSPSQESPNAPTSTAVLGSWSGPQPSLAPRBQEAPGVEPEWLDETSGIYYTE
                     IREREREKSEKGRPPWAAAVGTPPQARTSLQPHHHPWRPSVRESLCSTWPWNEDFNQ
                     KFTQLLLLQRSHPRSQDPLVKRSWPDVEENRGHLIEIRDLFGPGLDTQEPRIVILQG
                     AAGIGKSTLARQVREAWGRGQLYGRDQHFVFVFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQADALLGSLLGKTLLP
                     EASPLITARTTALQNLIPSLEQARWEVLGFSSESSRKEYFYRYFTDEROAIFRLVK
                     SNKELWALCLVPFVSWSWLACTCLMQQMRKEKLTLTSKTTTTLCLIHYLAQALQAOPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHIPLSYSFIHL
                     CFQEFPAAMSYYLEDEKGRGKHSNCIIDLEKTLEAYGIHGHLFGASTRFLGLLSDEG
                     EREMENIPHCRLSQGRNLMOWPSLQLLLQPHSLESLHCLYETRNKFLTQVMAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKLQLIEGROHRSTWSPTMVVLFRVFVPVDAYWQ
                     ILFSVLKVTRNLKELDLSFNVLITDAGAKHLCQRLROPSCKLQRLQLVSCGLTSDCCQDL
                     LAFGLRANQTLTELDLSFNVLITDAGAKHLCQRLROPSCKLQRLQLVSCGLTSDCCQDL
                     ASVLSASPSLKRLDLQQNNLDDVGVRLLCEGLRHPACHLIRLGLDQTTLSDEMRQELR
                     ALEQEKPQLLIFSRRKPSVMPTEGLDTGEMSNSTSSLKRQRLGSSERAASHVAQANLK
                     LLDVSKIFPIAEIAAGKSHEESSPEVVFVELLCVPSPASQGLDLHTKRQLGTDDDPWGPTG
                     PVATEVVDKEKNLYRVHFPVAGSYRWNTGLCPVMREAVTVEIEFCWHDQFLGEINPQ
                     HSWMVAGPLLDIKABPGAVEAVHLPHFVALQGGHVDTPLSLFQWAHFKEEGMLLEKPARV
                     ELHHIVLENPSFSPLGVLLKMIHNALRFRIPIVTVSVLLVHFLRVETNFHLYLIPSDCS
                     IRKAIDDLEMKPQFVRIHKPPLTPLYWGCRYTVSVGSGSGMEILFRLKELELCYRSPGE
                     DQLFSEFYVGHLGSGIRLQVRDKDETLVWEALVKPGRNTSQPWNLRCNRDARRY"
BASE COUNT     1247 a   1497 c   1454 g   1108 t
ORIGIN
```

```
Query Match             78.7%;  Score 4281.8;  DB 9;  Length 5306;
Best Local Similarity   96.8%;
Matches 4447;  Conservative    0;  Mismatches    2;  Indels  144;  Gaps   2;

Qy      1  GCCCCAGGGCCTGAGGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCGCCGGGGCTCACTCT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     31  GCCCCAGGGCCCTGAGGAGGTCTGAAGAAAACCTGGGAGCCAGCAGCGCCGGGGCTCCACTCT  90

Qy     61  GGGGTTCTGAAAGCCCCATTCCCCGCTCTCGCGGCTCCTCCCGACCCCCACACCCCTTCAGCCTT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     91  GGGGTTCTGAAAGCCCCATTCCCCTGCTTCGCGGCTCCTCCCGACCCCCACCCCTCTTCAGCCTT  150

Qy    121  GCAGCTCAAGGGTTGGATCTCAGGAGTCTGAAGAAAACCTGGGAGCCAGCAGGAATCTGAGGAACA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151  GCAGCTCAAGGGTTGGATCTCAGGAGTCTGAAGAAAACCTGGGAGCCAGCAGGAGGAATCTGAGGAACA  210

Qy    181  CAGAACAGTGAGCGTTGCCCACACCCCACCTCCCGTCACCAGGAGAGGGAAGAATCTGAGGAACA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211  CAGAACAGTGAGCGTTGCCCACACCCCATCTCCCGTCACCAGGAGAGGGAAGAATCTGAGGAACA  270

Qy    241  CCTCCCTGCCTGGCCCTGGACCCCCCATCCCAGGAGACGCTCCTCCCTATCAGCTGACTTCTTCCAGT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271  CCTCCCTGCCTGGCCCTGGACCCCCCATCCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT  330
```

```
Qy    301  GTCTTGCAGGCCCCTCTGGGGCTTCCTCCCCCTGGCTTTTTCCTACCACTCCCCTCTAT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    331  GTCTTGCAGGCCCCTCTGGGGCTTTTTCCTACCACTCCCCTCTAT                390

Qy    361  CGGCCGGTCTATCTGTAGGTGCCCTGGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391  CGGCCGGTCTATCTGTAGGTGCCCTGGGGATTTATAAAACTGGGTTCCGAATGCTGAATAAGA  450

Qy    421  GACGGGTAAGAGCCCAAGGCAAAGGACAGCAGCACTGTTTCTCTGCCCTGCCTGATACCCTCACCAC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451  GACGGGTAAGAGCCCAAGGCAAAGGACAGCAGCACTGTTTCTCTGCCCTGCCTGATACCCTCACCAC  510

Qy    481  CTGGGAAACATCCCCCAGACACACCCTTCTTAACTCCCGGGACACAGAGATGGCTGGCGGACCCTGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    511  CTGGGAAACATCCCCCAGACACACCCTTCTTAACTCCCGGGACACAGAGATGGCTGGCGGACCCTGG  570

Qy    541  GGCCGGCCTTGGGCCTTGTTACTTGGAGTTCCTGAAGAAGGAGGAGGAGTTCCAGCTT  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    571  GGCCGGCCTTGGGCCTTGTTACTTGGAGTTCCTGAAGAAGGAGGAGGAGTTCCAGCTT  630

Qy    601  CTGCTTCGCCAATAAAAGCGGCACTCCAGGAGGAGCTCTTCGGGGTGGTGAGACACCCGCCTCAGCCAGAG  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    631  CTGCTTCGCCAATAAAAGCGGCACTCCAGGAGGAGCTCTTCGGGGTGGTGAGACACCCGCCTCAGCCAGAG  690

Qy    661  AAGAGGAGTGGGCATGGAGGTGGCCTCGTACCTGGTTGGGCTCAGTGATGGGGAGGCACGGGGCC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    691  AAGAGGAGTGGGCATGGAGGTGGCCTCGTACCTGGTTGGGCTCAGTGATGGGGAGGCACGGGGCC  750

Qy    721  TGGGGACCTTAGCGCCTTCCATAACCTGGGAGCAGATGGGGGCTGAGGTCACTGTCCGCCCAAGCC  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    751  TGGGGACCTTAGCGCCTTCCATAACCTGGGAGCAGATGGGGGCTGAGGTCACTGTCCGCCCAAGCC  810

Qy    781  CAGGAAGGGGGCAGGCCACTCTCCCTCATTCCCCTACAGCCCCAAGTGAACCCCACCTGGGG  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    811  CAGGAAGGGGGCAGGCCACTCTCCCTCATTCCCCTACAGCCCCAAGTGAACCCCACCTGGGG  870

Qy    841  TCTCCCAGCCAACCCACCTCCACCGCAGGTGCTAATGCCCTGGATCCATGAATTGCCGCGG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    871  TCTCCCAGCCAACCCACCTCCACCGCAGGTGCTAATGCCCTGGATCCATGAATTGCCGCG  930

Qy    901  GGGTGCACCCAGGGGCTCAGAGGAGAAGGGGTTTTGAGACACTGCCTGACCATCTGGACCGC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    931  GGGTGCACCCAGGGGCTCAGAGGAGAAGGGGTTTTGAGACACTGCCTGACCATCTGGACC  990

Qy    961  CGCTGGAGAGAAATCTCTGCCTCACTCCTCTCCAAGCTTCTTCCAAGCTCCCCAGACCAT  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    991  CGCTGGAGAGAAAATCTCTGCCTCACTCCTTACCAAGCTTCTTCCAAGCTCCCCAGACCAT  1050

Qy   1021  GAGGTCTCCAAGCCCAGGAGTCACCCACGCCCCACATCCACAGCAGTGCTGCTGGGGGCTGG  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1051  GAGGTCTCCAAGCCCAGGAGTCACCCAAGCGCCCACATCCACAGTGCTGGGGGAGCTGG  1110

Qy   1081  GGATCCCCACCTCAGCCCTAGCACCCAGCGCCCAGCCCATGGGGACCCCAATGG  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1111  GGATCCCCACCTCAGCCCTAGCACCCAGCGCCCAGCCCATGGGGACCCCAATGG  1170

Qy   1141  CCTCTGGATGAAACGTCAGGAATTTACTACACAGAAAATCAGAGAAAGAGAGGAGAGAAA  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171  CCTTCTGGATGAAACGTCAGGAATTTACTACACAGAAAATCAGAGAAAGAGAGGAGAGAAA  1230

Qy   1201  TCAGAGAAAAGGCAGGCCCCCCATGGGCAGCGGCGGTGGTAGGGAACGCGCCGCACACC  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231  TCAGAGAAAAGGCAGGCCCCCCATGGGCGCGGTGGTAGGGAACGCGCCGCACACC  1290

Qy   1261  AGCCTACAGCCCCACCACACCCCCCCATGGGAGCCTTCTGTGTGAGGACCTTCGTTTCCACA  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1291  AGCCTACAGCCCCACCACACCCCCCCATGGGAGCCTTCTGTGTGAGGAGCCCTTCGTTTCCACA  1350

Qy   1321  TGGCCCCTGGAAAAATGAGGAATTTTAACCAAAAATTCACAGCCTGCTCTACTTCTACAAAGA  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351  TGGCCCCTGGAAAAATGAGGAATTTTAACCAAAAATTCACAGCCTGCTCTACTTCTACAAAGA  1410

Qy   1381  CCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGAAGCTGGCCTTGAATTATGTGGAGGAG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411  TCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGAAGCTGGCCTTGAATTATGTGGAGGAG  1470

Qy   1441  AATCGGAGGACATTTAATTCAGAGACTTATTTGGCCCAGGCCTTGGATACCCAAGAA  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471  AATCGGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAGGCCTTGGATACCCAAGAA  1530

Qy   1501  CCTCCGCATAGTCATCATCTGCAGGGGGGCTGCTCGGGAAATTGGGGAAGTCAACACTGGCCCAGGCAG  1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1531  CCTCCGCATAGTCATACTGCAGGGGGTGCTCGGGAAATTGGGGAAGTCAACACTGGCCCAGGCAG  1590

Qy   1561  GTGAAGGAAGCCTGGGGGGAGAGGCCAGCTGTATGGGGGGACCGCTTCCAGCCATGTTCTCTAC  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1591  GTGAAGGAAGCCTGGGGGAGAGGCCAGCTGTATGGGGGACCGCTTCCAGCCATGTTCTCTAC  1650

Qy   1621  TTCAGCTGCAGAGAGCTGGGCCCAGTCCAAGGTGGTGAGTCTCCGCTGAGCTCTC  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1651  TTCAGCTGCAGAGAGCTGGGCCCAGTCCAAGGTGGTGAGTCTCCGCTGAGCTCTC  1710

Qy   1681  GATGGGACAGCCACTCCCGGCTCCCATTAGACAGATCCTGTCTGCTAGGCCCAGGGCTGCTC  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1711  GATGGGACAGCCACTCCCGGCTCCCATTAGACAGATCCTGTCTGCTAGGCGGGCTGCTC  1770

Qy   1741  TTCATCCTCGATGGGTAGTAGATGAGCCAGGATGGGGTCTTGCAGAGGCGCAGTTCTGAGCTC  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1771  TTCATCCTCGATGGTTGGTAGATGAGCCAGGATGGGGTCTTGCAGAGGCCGAGTTCTGAGCTC  1830

Qy   1801  TGTCTGCACTGGAGCCCAGCCACAGCCGCGGCGATGGCACTGCTGGGGCAGTTTGCTGGGGAAA  1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1831  TGTCTGCACTGGAGCCCAGCCACAGCCGCGGGATGGCACTGCTGGGGCAGTTTGCTGGGGAAA  1890

Qy   1861  ACTATACTTCCCGGAGGCATCCTTCCTGATCACGGCCACAGCTCTGCAGGAACCTC  1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1891  ACTATACTTCCCGAGGCATCCTTCCTGATCACGGCCACAGCTCTGCAGGAACCTC  1950

Qy   1921  ATTCCTTCTTGGAGCAGCGTTGGGGTAGAGGTTCTCGAGTCCAGCAGG  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1951  ATTCCTTCTTTGGGAGCAGCACGTTGGGGTAGAGGTTCTCGAGTCCAGCAGG  2010

Qy   1981  AAGGAATATATTCTACAGATAATTCACAGATAATTAGAGCCAAGCAAATTGAGGTTTG  2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2011  AAGGAATATATTTCTACAGATAATTCACAGATGAAAAGGCAAAGCAAATTTAGGTTTG  2070

Qy   2041  GTCAAAATCAAAACAAAAGAGAGCTCTCGGGCCCTGTTCTTGTCCCTGGGTGTCTGGGGGCC  2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2071  GTCAAAATCAAAACAAAAGAGCTCTCGGGCCCGTGTTCTTGTCCCTGGGTGTCTGGGGCC  2130

Qy   2101  TGCACTTGCCTGATCGCAGCAGATGAAGCGGGAAGGAAAAACTCACACTGACTTCCAAGACC  2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2131  TGCACTTGCCTGATCGCAGCAGATGAAGCGGGAAGGAAAAACTCACACTGACTTCCAAGACC  2190

Qy   2161  ACCACAACCCCTCGTCTACATTACCCTTGCCCAGGCCTCAGCCATTCCAAGCTCAGCCATTGGGGACCCC  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2191  ACCACAACCCCTCGTCTACATTACCTTGCCCAGGCCTCAGCCATTCCAAGCTCAGCCATTGGGGACCC  2250

Qy   2221  CAGCTCAGAGACCTCTGCTCTGCTCTGTGGGCCAAAAAAGACCCTTTTC  2280
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2251  CAGCTCAGAGACCTCTGCTCTGTGGGCCATCTGGCAAAAAAGACCCTTTTC  2310

Qy   2281  AGTCCAGATGACCTCAGGAAGCAATGGGTTAGATGGGGCCATCACTCCCACCTTCTTGAAG  2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2311  AGTCCAGATGACCTCAGGAAGCAATGGGTTAGATGGGGCCATCACTCCCACCTTCTTGAAG  2370

Qy   2341  ATGGGTATTCTTCAAGAGCACCCCATCCCTTCATTCACCTGTTTC  2400
           |||||||||||||||||||||||||||||||||||||||||||||
Db   2371  ATGGGTATTCTTCAAGAGCACCCCATCCCTTCATTCACCTCTGTTTC  2430

Qy   2401  CAAGAGTTCTTTGCAGCAATGTCCTTATGTCTTGGAGGATGAGAGAGGAGGGTAAACAT  2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2431  CAAGAGTTCTTTGCAGCAATGTCCTTATGTCTTGGAGGATGAGAGAGGAGGGTAAACAT  2490

Qy   2461  TCTAATTGCATCAATAGATTTGGAAAAGACGCTAGAAGCAATATGGAATACATGGGCCTGTTT  2520
```

RESULT 8
AX089761

```
LOCUS       AX089761                4422 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 1 from Patent WO0116170.
ACCESSION   AX089761
VERSION     AX089761.1  GI:13443933
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,J.C.
  TITLE     Card proteins involved in cell death regulation
  JOURNAL   Patent: WO 0116170-A 1 08-MAR-2001;
            The Burnham Institute (US)
FEATURES             Location/Qualifiers
     source          1..4422
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..4422
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC34899.1"
                     /db_xref="GI:13443934"
                     /translation="MAGGAWGRLACYLEFLIKEBELKEFQLLLANKAHSRSSGETPAQ
                     PEKTSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQAQBGAGHSPSFPYSPSE
                     PHLGGSPQPTSTAVLMPWIHELPAGTQGSERRVLRQLPDTSGSRRWREISASLLYQAL
                     PSSPDHHSPSQESPNAPTSTAVLGWSSPPQPSLARREQEAPGTQWPLDETSGIYYTE
                     IREREKSEKGRPPWAAAVGTPPQOAHTSLQPHHHPWEPSVRESLCSTTWPWKNEDFNQ
                     KFTQLLLLQRPHPRSQDPLVKRSWPDYVEERNGHLIEIRDLFGPGLDTQEPRIVILOQG
                     AAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILP
                     EASFLITARTTALQNLIPSLEQARWRVLGFSESSRKEYFYRYFTDEQAIRAFRLVK
                     SNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTILCLHYLAQALDAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDLRKHGLDGAIISTFLKMGILQEHPIPLSYSFIHL
                     EREMENIFHCRLSQGRNLMQWVPSLQLLLOQPHSLESLHCLYETRNKIFLTQVMAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWSPTMVLFRWVPVTDAYWQ
                     ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDCKD
                     LAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLORLQLVSCGLTSDCCCDL
                     ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMROELR
                     ALEQEKPQLLIPSRRKPSVMTPTEGLDTGEMSNSTSLSKRQRLGSERAASHVAQANLK
                     LLDVSKIFPIAEBESSPEVVPEVELLCVPSPASQGDLHTKPLGTDDDFWGPTGPVAT
                     EVDKEKNLVRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGEINPQHSWM
                     VAGPLDIKAEPGAVEAVHLPHFVALQGQDVDTSLFQMAHFKEBGMLLEKPARVELHH
                     IVLENPSFSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIIPSDCSIRKA
                     IDDLEMKFQFVRIHKPPLTPLYMGCRYTVSGSGSGMLEILPKELELCYRSPGEDQLF
                     SEFYVGHLGSGIRLQVRDKDKDETLVWEALVKPGDLMEPVILTPPARIAVPSPLDAPQL
                     LHFVDQYREQLIARVISVEVVLDKLHGQVLSQEQYERVLAENTRPSQMRKLFSLSQSW
                     DRKCKDGLYQALKETHPHLIMELWEKGSKKGLLPLSS"
BASE COUNT    1043 a   1224 c   1240 g    915 t
ORIGIN
```

```
Qy          Best Local Similarity   76.2%;
            Query Match   97.0%;   Score 4148;   DB 6;   Length 4422;
            Matches 4290; Conservative   0; Mismatches   0; Indels 132; Gaps  1;
```

```
Qy    523  ATGGCTGGGGGAGGCGGCGCCGGCCGCTGGGGTCGTC...        582
           |||||||||||||||||||||||||||||||||||||||
Db      1  ATGGCTGGGGGAGGCGGCGCCGGCCGCTGGGGTCGTC...         60

Qy    583  CTGAAGGAGAGTTCCAGCTTCGTCTCGCCAATAAAGCG...       642
           |||||||||||||||||||||||||||||||||||||||
Db     61  CTGAAGGAGTTCCAGCTTCTGCTCGCCAATAAACGCA...        120

Qy    643  ACACCCGCTCAGCGAGGAAGAGCAGTGGCATGGCATG...       702
           |||||||||||||||||||||||||||||||||||||||
Db    121  ACACCCGCTCAGCCAGAGAAGAGCAGTGGCATGGCATG...       180

Qy    703  TATGGGGGAGCAGCGGCGGCCGGCCCTGGGAGCAGATG...       762
           |||||||||||||||||||||||||||||||||||||||
Db    181  TATGGGGGAGCAGCGGCGGCCGGCCCTGGGAGCAGATG...       240
```

```
Qy  4063  CTGTTCCAAATGGCCCACTTAAAGAGGAGGAGGAGGGATGCTCCTGGAAGCCAGCCAGGGTG  4122
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  3541  CTGTTCCAAATGGCCCACTTAAAGAGGAGGATGGCTCCTGGAAGCCAGCCAGGGTG  3600

Qy  4123  GAGCTGCATCACATAGTTCTGGGAAAACCCCAGCTTCTCCCCTTGGGAGTCCTCCTGAAA  4182
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  GAGCTGCATCACATAGTTCTGGGAAAACCCCAGCTTCTCCCCTTGGGAGTCCTCCTGAAA  3660

Qy  4183  ATGATCCATTAAATGCCTGCGCTTCATTCCCGTCACCTCTGTGGTGTTTGCTTTACCACCGC  4242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  ATGATCCATTAAATGCCTGCGCTTCATTCCCGTCACCTCTGTGGTGTTTGCTTTACCACCGC  3720

Qy  4243  GTCCATCCTGAGGAAGTCACCTTCCACCTTCTACCTGACTCCAAGTGACTGCTCTCCATTCGG  4302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GTCCATCCTGAGGAAGTCACCTTCCACCTTCTACCTGATCCAAGTGACTGCTCTCCATTCGG  3780

Qy  4303  ------------------------------------------------------------  4303
Db  3781  AAGGCCATAGATGATCTAGAAAATGAAAATTCACACAAGCCACCCCCG  3840

Qy  4303  ------------------------------------------------------------  4302
Db  3841  CTGACCCCACTTTATATGGGCGTGTCGTTACACTGTGTCTGGGTCTGGTTCAGGGATGCTG  3900

Qy  4303  ------------------------------------------------------------  4302
Db  3901  GAAATACTTCCCCAAGGAACTGGAGCTTCGCTATCGTAGCCCTGGAGAAGAGACCAGCTGTTC  3960

Qy  3901  -----------AAGGAACTGGAGCTTCGCTATCGTAGCCCTGGAGAAGAGACCAGCTGTTC  3960

Qy  4351  TCCGGAGTTCTACGTTGGGCCACTTGGGCATCAGGGGATCAGGCCTGCGCTGCCGCGAGCAGGAGAA  4410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  TCCGGAGTTCTACGTTGGGCCACTTGGGCATCAGGGGATCAGGCCTGCGCTGCCGCGAGCAGGAGAA  4020

Qy  4411  GATGAGAGACTCTGGTGTGTGGGAGGGCCTTGGTGGTGAAAACCAGGAGAGATCTCATGCTCTGCCTGCAACTACT  4470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  GATGAGAGACTCTGGTGTGTGGGAGGGCCTTGGTGGTGAAAACCAGGAGAGATCTCATGCTCTGCCTGCAACTACT  4080

Qy  4471  CTGATCCCTCCAGCCCGCCATAGCCGTACCTTCACCTTCTGGCATTTCACCTCTGCTGCTCAC  4530
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  CTGATCCCTCCAGCCCGCCATAGCCGTACCTTCACCTTCTGGCATTTCACCTCTGCTGCTCAC  4140

Qy  4531  TTTGTGTGACCAGTATCGAGAGCACACTGATAGCCCGAGTGACATCGGGTGGAGGGTTGTCTTG  4590
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TTTGTGTGACCAGTATCGAGAGCACACTGATAGCCCGAGTGACATCGGGTGGAGGGTTGTCTTG  4200

Qy  4591  GACAAACTGCATGGACAGGTGCTCAGCCAGGAGCGAGTAGAGGAGGAGGGGCTGGCTGAGAAC  4650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  GACAAACTGCATGGACAGGTGCTCAGCCAGGAGCGAGTAGAGGAGGAGGGGCTGGCTGAGAAC  4260

Qy  4651  ACGAGGCCCAGCAGGATGCGGGAAGCTGTTCAGCTTGAGCCAGTGCTCTGGGACCGGAAGTGC  4710
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  ACGAGGCCCAGCAGGATGCGGGAAGCTGTTCAGCTTGAGCCAGTGCTCTGGGACCGGAAGTGC  4320

Qy  4711  AAAGATGGACTCTCACCAGCCCTCAAAGAGACCATCTCACTCTCATTATTGGAACTCTTGG  4770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  AAAGATGGACTCTCACCAGCCCTCAAAGAGACCATCTCACTCTCATTATTGGAACTCTTGG  4380

Qy  4771  GAGAAGGGCCAGCAAAAAGGGACTCCTGCCACTCAGCCAGCTGA  4812
          ||||||||||||||||||||||||||||||||||||||||||||
Db  4381  GAGAAGGGCCAGCAAAAAGGGACTCCTGCCACTCAGCCAGCTGA  4422
```

REFERENCE

```
LOCUS       AF229060                4422 bp    mRNA    linear   PRI 02-FEB-2001
DEFINITION  Homo sapiens NAC-beta splice variant (NAC) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF229060
VERSION     AF229060.1  GI:12656106
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4422)
  AUTHORS   Chu,Z.-L., Pio,F., Xie,Z., Godzik,A. and Reed,J.C.
  TITLE     NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c
            pathway for apoptosis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4422)
  AUTHORS   Chu,Z.-L. and Reed,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The
            Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA
            92037, USA
FEATURES             Location/Qualifiers
     source          1..4422
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /cell_type="activated Jurkat T cells"
     gene            1..4422
                     /gene="NAC"
     CDS             1..4422
                     /gene="NAC"
                     /note="Apaf-1/Ced-4 family apoptosis regulator; contains
                     NB-ARC domain, CARD domain, and leucine-rich repeats
                     (LRR)"
                     /codon_start=1
                     /product="NAC-beta splice variant"
                     /protein_id="AAK00749.1"
                     /db_xref="GI:12656107"
                     /translation="MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGSTPAQ
                     PEKTSGMEVASYLVAQYGQRAWDLALHTWEBQMGLRSLCAQGEGAGHSPSFPYSPSE
                     PHLGSPSOPTSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRRWRRIISASLLYQAL
                     PSSPDHESPRSQESPNAPRTSTAVLGSWGSPPQPSLAPREGEAPGTQWPLDETSGIYYTE
                     IREREREKSEKGRRPWAAVVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDPNQ
                     KFTQLLLLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFGPGLDTOEPRIVILQG
                     AAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWQPQPADALLGSLLGKTILP
                     EASFLITARTTALQNLIPSLEQARWEVLGFSSSSRKEYFYRFTDEQOAIRAPRLVK
                     SNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTLCLHYLAQALOAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDLRKHGLDGAIISTFLKMGILQEHFIPLSISFIHL
                     CQFEFFAAMSYVLEDEKGRGKSNCIIDTLEAYGHGLGFGASTTRFLLGLLSDEG
                     ERMENIEHCRLSOGRNLMOWVPSLQLLLQPHSLESLHCLYETRNKTFLTQVWAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVLFRWVPVTDAYWQ
                     ILFSVLKYTRNLKELDLSGNSLSHSAVKSLCKTLRPRCLLETLRLACCGLTAEDCKD
                     LAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLTSDCCQDL
                     ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTLSDEMRQELR
                     ALEQEKPQLLLIFSBRKRSVMTFPVEGLCDPGMXNNSTSSLKRQRLGSSERRASHVAQANLK
                     LLDVSKIFPIAEIAEESSPEVVPVELLCVPSPASQGDLHTKPLGTDDDFWGPTGPVAT
                     EVDKEKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIBFCVWDOFGIEINFQHSWM
                     VAGPLLDIKAEPGAVEAVHLPHFVALCQGHVDTSLFQWAHFKEEGMLLEKPARVELHH
                     IVLENPSFSSPLGVLLKMIHNALRFIPVTSVVLLYHRVHPEEVTFHLYLIPSDCSIRKA
                     IDDLEMKROFVRIHKPPLTPLTWGCRTVSGSGSGMLEILPKLELELCYRSPGEDQLF
                     SEYVGHLGSGIRLQVDKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQL
                     LHFVDQYFBQLIARVTSVEVVLDKIHGQVLSOEYERVLAENTRPSQMRKLFSLSQSW
                     DRCKDGIYQALKETHPHLIMELWEKGSKKGLIPLSS"
BASE COUNT     1043 a   1224 c   1240 g    915 t
ORIGIN
```

```
RESULT 9
AF229060

Query Match           76.2%;  Score 4148;  DB 9;  Length 4422;
Best Local Similarity 97.0%;
Matches 4290;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps 1;

Qy  523  ATGGCTGGGCGGGCTGGGCGGCCGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG  582
         |||||||||||||||| ||||||||||| |||||| |||||||||||||||||||||||||
Db    1  ATGGCTGGCGGGCCGGAGCCTGGGCCGGGCCTGCC                           60

Qy  583  CTGAAGGAGTTCCAGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTAGCAG  642
         ||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||||
Db   61  CTGAAGGAGTTCCAGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG  120

Qy  643  ACACCCGCTCAGCCAGAGAAGACAGAGTGGCCATGGAGGTGGCCTCGTACCTGGGTGGCTCAG  702
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

[DNA sequence alignment, rotated 90°, Qy/Db rows with base-pair coordinates from ~2341 to ~4812]

RESULT 10
AF229061
LOCUS   AF229061   4200 bp   mRNA   linear   PRI 02-FEB-2001
DEFINITION   Homo sapiens NAC-gamma splice variant (NAC) mRNA, complete cds, alternatively spliced.

```
ACCESSION   AF229061
VERSION     AF229061.1  GI:12656108
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4200)
  AUTHORS   Chu,Z.-L., Pio,F., Xie,Z., Godzik,A. and Reed,J.C.
  TITLE     NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c
            pathway for apoptosis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4200)
  AUTHORS   Chu,Z.-L. and Reed,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The
            Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA
            92037, USA
FEATURES             Location/Qualifiers
     source          1..4200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /cell_type="activated Jurkat T cells"
     gene            1..4200
                     /gene="NAC"
     CDS             1..4200
                     /gene="NAC"
                     /note="Apaf-1/ced-4 family apoptosis regulator; contains
                     NB-ARC domain, CARD domain, and leucine-rich repeats
                     (LRR)"
                     /codon_start=1
                     /product="NAC-gamma splice variant"
                     /protein_id="AAK00750.1"
                     /db_xref="GI:12656109"
                     /translation="MAGGAWGRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQ
                     PEKTSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQOEBGAGHSPSFPYSPSE
                     PHLGSPSGPTSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRRWRBISASLLYQAL
                     PSSPDNEHSPSQOEESPNAPTSTAVLGSWGSPQOPSLAPREQEARGTOWPLDETSGIYYTE
                     IRERERKEKSEKGRPWAAAVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQ
                     KFTQLLLLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFPGLDTQBPRIVILQG
                     AAGIGKSTLARQVKEAWGRQLYGDRFQHVPYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQOPQPADALLGSLLGKTILP
                     EASFLITARTTALQNIPSLEQARWVEVLGFSSBSSRKYEYFYRIPDERQAIRARVLVK
                     SNKEIWALCLVPWVSWIACTCLMQOMRKREKTLLTSKTTTLCIHYLAQOALOAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIIISTFLKMGILQEHPIPLSYSFTHL
                     CFQEFPAAMSTVLEDEKGRGKHSNCIIIDLKEXTLEAYGIHGLFGASTTVLFGLLSDEG
                     EREMENIFHCRLSQGRNLMOWVPSLQLLLQPHSLSHLCLVETRNKTFLTQVMAHFEE
                     MGMCVDTDMELLVCTFCIKFSSHVKKLQLIEGRQHRSTWSPTMVVLFRWVFVTDAYWQ
                     ILFSVLKVTRNLKEDLLSGNSLSHAVKSLCKTLRRPRCLLETTLRLAGCGLTAEDCKD
                     LAPGIRANQTLTELDLSFNVLTDAGAKHLCORLROPSCKLQRLQLVSCGLTSDCCODL
                     ASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLLRIGKFPSVMTPTEGLDTGE
                     MSNSTSSLKRORLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSPEVVPBELLCVP
                     SPASOGDLHTTKPLGTDDDFWGPTGVPVATEVVDKEKNLYRVHFPVAGSYRWNTGLCFV
                     MREAVTVEIEFCVWDQPLGEINPQHSWMVAGPLLDIKAEPGAVEAVLHPHVALQGGH
                     VDTSLFOMAHFKEBGMLLEKPARVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSV
                     VLLYHRVHPEEVTFHLYLIPSDCSIRKELELCYRBPGEDOLFSEFYVGHLGSSGIRLQV
                     KDKKDETLVHEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLHFVDQYREQLIARVT
                     SVEVVLDKLHQQVLSOEQYERVLAENTRPSQMRKLFSLSQSWDRKCKDGLYQALKETH
                     PHLIMELWEKGSKKGLLPLSS"
BASE COUNT      986 a   1167 c   1180 g    867 t
ORIGIN
```

```
Query Match   75.3%;  Score 4100;  DB 9;  Length 4200;
Best Local Similarity   97.9%;  Pred. No. 0;
Matches 4200;  Conservative  0;  Mismatches  0;  Indels 90;  Gaps 1;

Qy    523 ATGGCTGGCGGAGCCTGGGGGCGGCCGGCCTGGGCCTGGTTGTTACTTGGGAGTTCCTGAA 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGCGGAGCCTGGGGGCGCCTGGCCTGCTACCTGGAGTTCCTGAAGAAGGAGGAG 60

Qy    583 CTGAAGGAGTTCCAGCTTCTGCTCGCGCCAATAAAGCGCACTCTCAGGAGCTCTTCGGGGTGAG 642
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGAAGGAGTTCCAGCTTCTGCTCGCGCCAATAAAGCGCACTCTCAGGAGCTCTTCGGGGTGAG 120

Qy    643 ACACCCGCTCAGCCAGAGAAGACGGAGTGGCCATGGAGGTGGCCTCGTTACCTGGGCTCAG 702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACACCCGCTCAGCCAGAGAAGACGGAGTGGCCATGGAGGTGGCCTCGTTACCTGGGCTCAG 180

Qy    703 TATGGGGGAGCAGCGGGCGGCCTGGGACCTAGCCCTTCCATACCTGGGGAGCAGATGGGGCTGAGG 762
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATGGGGGAGCAGCGGGCGGCCTGGGACCTAGCCCTTCCATACCTGGGGAGCAGATGGGGCTGAGG 240

Qy    763 TCACTGTCGCCAAGCCCAGAGAAGGAGGCAGGCCAGCTCACTCTTCCCCTCATTCCCCTACAGCCCA 822
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTGTCGCCAAGCCCAGAGAAGGAGGCAGGCCAGCTCACTCTTCCCCTCATTCCCCTACAGCCCA 300

Qy    823 AGTGAACCCCACCTGGGTCTCCCAGCCAACCCAACCTCCACCTCCCCTCCCCTACAGCCCA 882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGTGAACCCCACCTGGGTCTCCCAGCCAACCCAACCTCCACCTCCCCTCCCCTACAGCCCA 360

Qy    883 ATCCATGAATTGCCGGCGGGTGCAGCCCAGGGGCTCAGAGAGGAAGGGGTTTTGAGGACAGTG 942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCCATGAATTGCCGGCGGGTGCAGCCCAGGGGCTCAGAGAGGAAGGGGTTTTGAGGACAGTG 420

Qy    943 CCTGACACATCTGGACGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCTTACCAAGCTCTT 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTGACACATCTGGACGCGCGCTGGAGAGAGAAATCTCTGCCTCACTCCTCTTACCAAGCTCTT 480

Qy   1003 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAGGCCAAGCGCCCCCACATCCACA 1062
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAGGCCAAGCGCCCCCACATCCACA 540

Qy   1063 GCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCAGCCTAGCACCCAGAGAGGCAGGAG 1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCAGCCTAGCACCCAGAGAGGCAGGAG 600

Qy   1123 GCTCCCTGGGACCCAAATGGCCTCTCGGATGAAAACGTTCAGAAATTTACTACAGAAATCAGA 1182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCCTGGGACCCAAATGGCCTCTCGGATGAAAACGTTCAGAAATTTACTACAGAAATCAGA 660

Qy   1183 GAAAGAGAGAGAGAAATCAGAGAAAAGGCAGGCGCCCCCATGGGCAGCGGGTGGTAGGAACG 1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GAAAGAGAGAGAGAAATCAGAGAAAAGGCAGGCGCCCCCATGGGCAGCGGGTGGTAGGAACG 720

Qy   1243 CCCCCACAGGCGCACACCAGCCTACAGCCCCACCCACCCACCATGGGAGCCTTCTGTGAGA 1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCCCACAGGCGCACACCAGCCTACAGCCCCACCCACCCACCATGGGAGCCTTCTGTGAGA 780

Qy   1303 GAGAGCCCTCTGTTCCACATGGGCCCTGAAAAAATGAGGATTTAACCAAAAAATTCACACAG 1362
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGAGCCCTCTGTTCCACATGGGCCCTGAAAAAATGAGGATTTAACCAAAAAATTCACACAG 840

Qy   1363 CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGG 1422
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGG 900

Qy   1423 CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 1482
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 960

Qy   1483 GGGCTGGATACCCAAGAACCTCGCATAGTCATACTGCAGGGGGCTGCTCGGAATTGGGGAG 1542
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGGCTGGATACCCAAGAACCTCGCATAGTCATACTGCAGGGGGCTGCTCGGAATTGGGGAG 1020

Qy   1543 TCAACACTGGCCAGGCAGTGAAGGAAGCCTGGGGACGGCAGCTGTATGGGGGACCGC 1602
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCAACACTGGCCAGGCAGTGAAGGAAGCCTGGGGACGGCAGCTGTATGGGGGACCGC 1080

Qy   1603 TTCCAGCATGTCTTCTACTTCAGTGCAGAGGAGCTGGCCCAGTCGAGAGGTTGGGTGAGTCTC 1662
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTCCAGCATGTCTTCTACTTCAGTGCAGAGGAGCTGGCCCAGTCGAGAGGTTGGGTGAGTCTC 1140

Qy   1663 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGTTCCCAATTAGACAGATCCTGTCT 1722
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTCCAGCATGTCTTCTACTTCAGTGCAGAGGAGCTGGCCCAGTCGAGAGGTTGGGTGAGTCTC 1140
```

```
Db   1141 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCT 1200

Qy   1723 AGGCCAGAGCGGCTGCTCTTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCAG 1782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGGCCAGAGCGGCTGCTCTTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCAG 1260

Qy   1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG 1842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG 1320

Qy   1843 GGCAGTTTGCTGGGGAAAACTATACTTCCCGAGGCATCCTTCCTGATCACGGCTCGGACC 1902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGCAGTTTGCTGGGGAAAACTATACTTCCCGAGGCATCCTTCCTGATCACGGCTCGGACC 1380

Qy   1903 ACAGCTCTGCAGAACCTCATTCCTTCTTTGGAGCAGGCACGTTGGGTAGAGGTCCTGGGG 1962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACAGCTCTGCAGAACCTCATTCCTTCTTTGGAGCAGGCACGTTGGGTAGAGGTCCTGGGG 1440

Qy   1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCA 2022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTCACAGATGAAAGGCAAGCA 1500

Qy   2023 ATTAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCCC 2082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATTAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCCC 1560

Qy   2083 TGGGTGTCCTGGCTGGCCTGCACTTGCCTGATGCAGCAGATGAAGCGGAAGGAAAAACTC 2142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TGGGTGTCCTGGCTGGCCTGCACTTGCCTGATGCAGCAGATGAAGCGGAAGGAAAAACTC 1620

Qy   2143 ACACTGACTTCCAAGACCACCACAACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAA 2202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ACACTGACTTCCAAGACCACCACAACCCTCTGTCTACATTACCTTGCCCAGGCTCTCCAA 1680

Qy   2203 GCTCAGCCATTGGGACCCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCTCAGCCATTGGGACCCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG 1740

Qy   2263 CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CAAAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC 1800

Qy   2323 ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860

Qy   2383 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGATGAG 2442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTCATTCACCTCTGTTTCCAAGAGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGATGAG 1920

Qy   2443 AAGGGGAGAGGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATAT 2502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGGGGAGAGGTAAACATTCTAATTGCATCATAGATTTGGAAAAGACGCTAGAAGCATAT 1980

Qy   2503 GGAATACATGGCCTGTTTGGGGCATCAACCACACGTTTCCTATTGGGCCTGTTAAGTGAT 2562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GGAATACATGGCCTGTTTGGGGCATCAACCACACGTTTCCTATTGGGCCTGTTAAGTGAT 2040

Qy   2563 GAGGGGGAGAGAGAGATGGAGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG 2622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GAGGGGGAGAGAGAGATGGAGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG 2100

Qy   2623 ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCTCCAC 2682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGGAGTCCCTCCAC 2160

Qy   2683 TGCTTGTACGAGACTCGGAACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAA 2742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGCTTGTACGAGACTCGGAACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAA 2220

Qy   2743 ATGGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCACTTTCTGCATTAAATTC 2802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCACTTTCTGCATTAAATTC 2280

Qy   2803 AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGCAGGCAGCACAGATCAACATGGAGC 2862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AGCCGCCACGTGAAGAAGCTTCAGCTGATTGAGGGCAGGCAGCACAGATCAACATGGAGC 2340

Qy   2863 CCCACCATGGTAGTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCCACCATGGTAGTCCTGTTCAGGTGGGTCCCAGTCACAGATGCCTATTGGCAGATTCTC 2400

Qy   2923 TTCTCCGTCCTCAAGGTCACCAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACTCGCTG 2982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTCTCCGTCCTCAAGGTCACCAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACTCGCTG 2460

Qy   2983 AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTCGCTGCCTCCTGGAG 3042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACGCCCTCGCTGCCTCCTGGAG 2520

Qy   3043 ACCCTGCGGTTGGCTGGCTGTGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 3102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACCCTGCGGTTGGCTGGCTGTGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 2580

Qy   3103 CTGAGAGCCAACCAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT 3162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CTGAGAGCCAACCAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT 2640

Qy   3163 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG 3222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG 2700

Qy   3223 CTGGTCAGCTGTGGCCTCACGTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCC 3282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CTGGTCAGCTGTGGCCTCACGTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCC 2760

Qy   3283 AGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAACAACCTGGATGACGTTGGCGTGCGA 3342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 AGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAACAACCTGGATGACGTTGGCGTGCGA 2820

Qy   3343 CTGCTCTGTGAGGGGCTCAGGCATCCTGCCTGCAAACTCATACGCCTGGGGCTGGACCAG 3402
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CTGCTCTGTGAGGGGCTCAGGCATCCTGCCTGCAAACTCATACGCCTGG----------- 2869

Qy   3403 ACAACTCTGAGTGATGAGATGAGGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG 3462

Db   2870 ------------------------------------------------------------ 2869

Qy   3463 CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG 3522
                             |||||||||||||||||||||||||||||||||||||||||
Db   2870 -------------------GGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG 2910

Qy   3523 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2911 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 2970

Qy   3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3030

Qy   3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCT 3702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3031 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCT 3090

Qy   3703 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3091 TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3150

Qy   3763 GGCCCCACGGGGCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTGTACCGAGTT 3822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3151 GGCCCCACGGGGCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTGTACCGAGTT 3210

Qy   3823 CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3211 CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3270
```

ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  AUTHORS Reed, J.C.
  TITLE   Card proteins involved in cell death regulation
  JOURNAL Patent: WO 011170-A 3 08-MAR-2001;
           The Burnham Institute (US)
FEATURES   Location/Qualifiers
     source   1..4194
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS   1..4194
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC34900.1"
                     /db_xref="GI:13443936"
                     /translation="MAGGAWRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQ
                     PEKITSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQQEQGAGHSPSPPYSPSE
                     PHLGSPSQPTSTAVLMPWIHELPAGTQGSRRVLRQLPDTSGRRWREISASLLYQAL
                     PSSPDHESPSPQSSPNAPTSTAVLGWSGPPQSLARREQEAPGTQWPLDETGSIYYTE
                     IREREREKSEKGRPPWAAAVGTPPQAHTSLQPHHHPMEPSVRESLCSTWPWKNEDFNQ
                     KFTQLILLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFGPLDDYDEPRIVILQQ
                     AAGIGKSTILARQVKEAWGRGQLYGDRRQFHVYFYFSCRELAQASKVVSLAELIGKDGIATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILP
                     SNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTLCLHYLAQALQAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDDLRKHGDGAIISTFLKMGILQEHPIPLSYSFIHL
                     CFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKYLEAYGIHGAIISTRFFLLGLLSDEG
                     EREMENIFHCRLSQGRNLMQWPSLQLLLQHSLESLHCLYETRNKFLTQVWMAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRWVPVTDAYWQ
                     ILFSVLKVTRNIKELDLSGNSISHSAVKSLCKTLRRPRCLLETTLRAELPSDCCODL
                     LAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQLVSCGLITSDCCODL
                     ASVLSASPSIKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLKPSWMTPTEGLDTGEM
                     SNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAIAEESSPEVVRELLCVPS
                     PASQGDLHTKPLGTDDDFWGPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVM
                     REAVTVEIBFCVWDQFLGEINPQHSMMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHV
                     DTSLFQMAHFKEEGMILEKPARVELHHIVLENPSFSPLGVLKMIHNALRFIPVTSVV
                     LLYHRVWHPEEVTFHLYLIFPSDCSIRELELCYRSPGEDQLFSEFYVGHLGSGIRLQVKD
                     KKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLHFVDQYREQIIARVTSV
                     EVVLDKLHGQVLSQBOQYERVLAENTRPSQMRKLFSLSQSWDRKCKDGLYQALKETHPH
                     LIMELWEKGSKKGLLPLSS"
BASE COUNT   984 a   1167 c   1176 g   867 t
ORIGIN

Query Match   74.9%; Score 4078; DB 6; Length 4194;
Best Local Similarity   97.8%;
Matches 4194; Conservative   0; Mismatches   0; Indels 96; Gaps   2;

Qy   3883 GAAGCGGTGACCGTTGAGATTGAATTCTGTGTGTTCCTTGGGGTGGGGTGGGGACCAGTCAAC 3942
Db   3271 GAAGCGGTGACCGTTGAGATTGAATTCTGTGTGTTCCTTGGGGTGGGGTGGGGACCAGTCAAC 3330

Qy   3943 CCACAGCACAGCTGGATAGGTGGCAGGGGCCTCTGCTCGCTGGACATCAAAGGCTGAGCCTGGAGCT 4002
Db   3331 CCACAGCACAGCTGGATGGTGGCAGGGGCCTCTGCTGCTCGCTGGACATCAAAGGCTGGAGCT 3390

Qy   4003 GTGGAAGCTGTGTGCACCTCCCTCACTTTGTGTGGGGGGCCCAGTGGACACATCC 4062
Db   3391 GTGGAAGCTGTGTGCACCTCCCTCACTTTGTGTGGGGGCCCAGTGGACACATCC 3450

Qy   4063 CTGTTCCAAATGCCCACTTTAAAAGAGGAGGGGGGTG 4122
Db   3451 CTGTTCAAATGCCCACTTTAAAAGAGGAGGGGGGTG 3510

Qy   4123 GAGCTGCATCACATAGTTCTGGAAAAACCCCAGCGTTCTCCTGCTTCCTGAAA 4182
Db   3511 GACCTGCATCACATAGTTCTGGAAAAACCCCAGCGTTCTCCTGCTTCCTGAAA 3570

Qy   4183 ATGATCCATAAATGCCCTGGGCTTCATTCCCGTGCACCTCTGTGTGGTTGCTTTACCACCGC 4242
Db   3571 ATGATCCATAAATGCCCTGGGCTTCATTCCCGTGCACCTCTGTGTGGTTGCTTTACCACCGC 3630

Qy   4243 GTCCATCCTGAGGAAGTCACCTTCCACCTTCCACCTTCACCTTCACTTCCAAGTGACTGCTCCATTCGG 4302
Db   3631 GTCCATCCTGAGGAAGTCACCTTCCACCTTCACCTTCACTTCCAAGTGACTGCTCCATTCGG 3690

Qy   4303 AAGGAACTGGAGCTCTGCTATCCGAAGCCCTGGAGGAAGACCAGCCAGTCCCATGTCTGAAA 4362
Db   3691 AAGGAACTGGAGCTCTGCTATCCGAAGCCCTGGAGGAAGACCAGCCAGTCCCATGTCTGAAA 3750

Qy   4363 GTTGGGCCACTTGGGATCAGGGGTGCAAGTGAGTGAAAGACAAGAACAAGAAAAGATGAGACTCTG 4422
Db   3751 GTTGGGCCACTTGGGATCAGGGGTGCAAGTGAGTGAAAGACAAGAACAAGAAAAGATGAGACTCTG 3810

Qy   4423 GTGTGGGAGGCCTTGGGATCAGGGGTGCAAGTGAGTGAAAGATGAGACTCTG 4482
Db   3811 GTGTGGGAGGCCCTTGGGATCAGGAAACCAGGAAGATCTGCTAGCACTACTCTGTATTCCCTCA 3870

Qy   4483 GCCCGCATAGCCGTACCTTCACCTCTCTGGATGCCCCGGCAGTTGCTGCTGCTGCTGGGACCAG 4542
Db   3871 GCCCGCATAGCCGTACCTTCACCTCTGGATGCCCCGGCAGTTGCTGCTGCTGCTGGGACCAG 3930

Qy   4543 TATCGGAGAGCAGCAGCTGAATAGCCCCAGTGACATCGGTGTGGAGGTTGTCTTCTTGGACAAAACTGGCAT 4602
Db   3931 TATCGGAGAGCAGCAGCTGAATAGCCCCAGTGACATCGGTGTGGAGGTTGTCTTCTTGGACAAAACTGCAT 3990

Qy   4603 GGACAGGTGCTGAGCCAGCAGCTACAGAGAGAGGTGTGGCTGAGAACACGCAGAGGAGGAGCCCAGC 4662
Db   3991 GGACAGGTGCTGAGCCAGCAGCTACAGAGAGAGGTGTGGCTGAGAACACGCAGAGGAGGAGCCCAGC 4050

Qy   4663 CAGATGCGGGAAGCTGTTCAGCTTGAGCGAGTCCTGGGACCGGAAGTGCAAAGATGGGACTC 4722
Db   4051 CAGATGCGGGAAGCTGTTCAGCTTGAGCGAGTCCTGGGACCGGAAGTGCAAAGATGGGACTC 4110

Qy   4723 TACCAAGCCCTGAAGGAGGAGGACCCATCCTCACCTCAGTCCTCGGGGACCAGCGGGAGGGGCAGC 4782
Db   4111 TACCAAGCCCTGAAGGAGGAGGACCCATCCTCACCTCAGTCCTCGGGGACCAGCGGGAGGGGCAGC 4782

Qy   4783 AAAAAGGGGACTCCTGCCACTCACCAGCTGA 4812
Db   4171 AAAAAGGGACTCCTGCCACTCACCAGCTGA 4200

Qy   523 ATGGCTGGCGGGAGCCCTGGGGGCCGCCGCCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 582
Db   1 ATGGCTGGCGGGAGCCCTGGGGGCCGCCGCCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG   60

Qy   583 CTGAAGGAGTTCCAGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
Db   61 CTGAAGGAGTTCCAGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120

Qy   643 ACACCCGCTCAGCCGCAGAAGAAGAGGCGAGTGGCATGGACATGGCCAGTGGCTTCGTTACCTGGTTCGGGCTCAG 702
Db   121 ACACCCGCTCAGCCGCAGAAGAAGAGGCGAGTGGCATGGACATGGCCAGTGGCTTCGTTACCTGGTTCGGGCTCAG 180

Qy   703 TATCGGGGAGCAGCGGGGGCCTGGGACCTAGCCCTTCCATACCTGGGAGGCAGAGATGGGGCTGAG 762
Db   181 TATGGGGAGCAGCGGGGGCCTGGGACCTAGCCCTTCCATACCTGGGAGGCAGAGATGGGGCTGAG 240

Qy   763 TCACTGTGCGCCCAAAGCCCCAGGAAGGGGCGAGGCCACTCCTCCATTCCCTCATTCCCTCACAGCCA 822
Db   241 TCACTGTGCGCCCAAAGCCCCAGGAAGGGGCGAGGCCACTCTCCATTCCCTCATTCCCTCACAGCCCA 300

Qy   823 AGTGAAACCCCACACTGGGGTCTTCCCAGCCAACCTCCACCGCAGTGCTGCTTAATGCCCTGG 882
Db   301 AGTGAAACCCCACACTGGGGTCTTCCCAGCCAACCTCCACCGCAGTGCTGCTTAATGCCCTGG 360

RESULT 11
LOCUS   AX089763   4194 bp   DNA   linear   PAT 21-MAR-2001
DEFINITION Sequence 3 from Patent WO0116170.
ACCESSION   AX089763
VERSION   AX089763.1 GI:13443935
KEYWORDS
SOURCE   Homo sapiens (human)

```
Qy    2521 ACCCTGCGCGGTTTGGCTGGGCTGGGCCTCGAGGACTGGCAAGGAACCTTGCCTTGGG 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ACCCTGCGCGGTTTGGCTGGGCTGGGCCTCGAGGACTGGCAAGGAACCTTGCCTTGGG 2580

Qy    3103 CTGGTCAGCTGTGTGGCCTCACGTCTGCCTCAGGACTGCCTCGTGTTAGTGTGCC 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CTGAGAGCCAACCAGACCCTGACCGGAGCTGGACTTCAATGTGCTCACGGATGCT 2640

Qy    2581 CTGAGAGCCAACCAGACCCTGACCGGAGCTGGACTTCAATGTGCTCACGGATGCT 2640
Db    3568 ATGATCCATAATGCCCTGCGCCTTCATTCCCGTCACCTGTGTGTTGTTTACCACGC 3627
```

RESULT 12

```
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /chromosome="11"
              /cell_type="activated Jurkat T cells"
gene          1..4332
              /gene="NAC"
CDS           1..4332
              /gene="NAC"
              /note="Apaf-1/Ced-4 family apoptosis regulator; contains
              NB-ARC domain, CARD domain, and leucine-rich repeats
              (LRR)"
              /codon_start=1
              /product="NAC-delta splice variant"
              /protein_id="AAK00751.1"
              /db_xref="GI:12656111"
              /translation="MAGGAWGRIACYLEFLKEELKEFQLLLANKAHSRSSSGEETPAQ
              PEKTSGMEVASYLVAQYGEQRAWDLALHTWEQMGLRSLCAQQEGAGHSPSFPYSPSE
              PHLGGPSSQPTSTAVLMPWIHELPAGTCQGSERRVLRQLPDTSGRRWREISASLLYQAL
              PSSPDHRSPSQESSPNAPTSTAVLGSWSGPPQPSLAPRREQEARGTQWPLDETSGIYYTE
              IRERERKSEKGRPPWAAVVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQ
              KFTQLLLQRPHPRSQDPLVKRSWPDTVEENRGHLIEIRDLFGPGIDTOEPRIVILOG
              AAGIGKSTLARQVKEAWRGQLYGDRPOHVFVFSCRELAQSKVVSLAELIGKDGTATP
              APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILP
              EASFLITARTTALQNLIPSLEQARWRVEVLGFSESSRKEYFRYFTDERQRAIRAFRLVK
              SNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTLCLHYLLAQALCQAQPLGP
              QLRDLCSLAAEGTWQKKTLFSPDDLRKHGLDGAIISTPLKMGILQEHPIPLSYSFIHL
              CFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAYGIHGLFGASTRFLLFGLLSDEG
              EREMENIFHCRLSQGRNLMQWPYSLQOLLIQEHPSELSHLCIYETRNKTFLTQVMAHFEE
              MGMCVEFDMELLVCTFCTKFSRHVKKLQLIEGRQHRSTWSPTMVVLRRWVPVTDAYNQ
              ILFSVLKVTRNLKELDLSGNSLSHSAVRSLCKTLRPRCLILETLRLAGCGLTAEDCKD
              LAFGLRANQTLTELDLSFNVLTDAGAKHLCQRLRQPSCKLORLVSCGLTSDCCQDL
              ASVLSASPSLKELDLOQNNLDDVGVRLLCEGLRHPACKLIRLGKPSVMTPIEGLDTGE
              SPASQGDLHTKPLGTDDDFWGPTGPVATEEVVDKERNLVRVHHPEVAGSYRWNTGLCFV
              MREAVTVEIEFCVWDQFLGEINPQHSWMVAGPLLDIKABPGAVEAVHLPHFVALQGGH
              VDTSLFQMAHFKEEGMLLEKPARVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSV
              VLIYHRVHPEEVTFHLYLIPSDCSIRKAIDDLEMKFQVRHHKPPLTPLYMGCRYTV
              SGSGSGMLEILPKELELCYRSPGEEDQLFSEFYVGHLGSSGIRLQVKDKKDETLVWEALV
              KPGDLMPATTLIPPARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVL
              SQEQYERVLAENTRPSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKK
              GLLPLSS"
ORIGIN
BASE COUNT     1018 a   1202 c   1211 g   901 t
```

Best Local Similarity   72.7%;
Matches 4200;   Conservative   95.0%;   Pred. No. 0;   Mismatches   0;   Indels   222;   Gaps   2;

Query Match   Score 3958;   DB 9;   Length 4332;

```
Qy      1 ATGGCTGGCGGGGAGGCCTGGGGCCGGATAGCCTGCTACCTGGAGTTCCTGAAGGAAGAGCTGAAGGAGTTTCAGCTG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGCGGGGAGGCCTGGGGCCGGATAGCCTGCTACCTGGAGTTCCTGAAGGAAGAGCTGAAGGAGTTTCAGCTG 60
```

```
RESULT 13
LOCUS       AX089765                4329 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 5 from Patent WO0116170.
ACCESSION   AX089765
VERSION     AX089765.1  GI:13443937
KEYWORDS    
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,J.C.
  TITLE     Card proteins involved in cell death regulation
  JOURNAL   Patent: WO 0116170-A 5 08-MAR-2001;
            The Burnham Institute (US)
FEATURES             Location/Qualifiers
     source          1..4329
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..4329
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC34901.1"
                     /db_xref="GI:13443938"
                     /translation="MAGGAWGRLACYLFLFKKEELKEFQLLLANKAHSRSSGETPAQ
                     PEKTSGMEVASYLVAQYGEQRAWDLALHTWBQMGLRSLCAQOEGAGHSPSFPYSPSE
                     PHLGSPSQPTSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQAL
                     PSSPDHESPSQESPNAPTSTPAVLSSWGSPPOPSLAPREQPAPGTOWPLDETSGIYYTE
                     IRERREKSEKGRPPWAAVVGTPPQQAHTSLQPHHHPWEPSVRESLCSTWPWKNDEFNQ
                     KFTQLLLLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLFGPGDPIRVIIQG
                     AAGIGKSTLARQVKEAWGRGQLYGDRFQEHVYFYFSCRELLAQSKVSLAELIGKDGTATP
                     APIRQILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKIILP
                     SNKELWALCLVPWSWLACTCLMOQMKREKLLTLISKTTTLLCLHYLAQLQAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILOGHPPLSYSFIHL
                     CPOEFFAAMSYVLEDEKGRGKHSNCIIIDLEKTLEAYGIHGLFGASTTRFLLGLSDEG
                     EREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLLESLLHCLYETRNKTFLTQVMAHFEE
                     MGMCVETDMELLVCTFCIKFSRHVKKLLQLIEGRQHRSTWSPTMVVLFRWPVTDAYWQ
                     ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTTAEDCKD
                     LAFGLRANQITLRELDLSFNVLTDAGAKHLCORLQPSCKLQRLQLVSCGLTSDCCODL
                     ASVLSASPSLKELDLQONNLDDVGVRLLCEGLRHPACKLIRLKPSVMTPTEGLDTGEM
                     SNSTSSLKRQRLGSERAASHVAQANLKLLDVSKIFPIAEIAEESSPEVVEVLLCVPS
                     PASQDLHTKPLGTDDFWGPTGPVATEVVDKEKNLYRVHFPPVAGSYRWNTIGLCFVM
                     REAVTVEIEFCVWDQFPLGEINPQHSWMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHV
                     DTSLPQMAHFKEEGMLLEKPARVELHHIVLENPSPSPLGVLLRKIMHNALRPIPVTSVV
                     LIYRVHPEEVTFHLYLILPSDCSTRKAIDDLEMKFQFQVRIHKPPLTPLYMGCRYTVS
                     GSGSGMLEILPKELELCYRSPGEDQLFSEFYVGHGLGSGIRLQVDKDKDETLVWEALVK
                     PGDLMPATLLIPPARIAVPSPLDAPOLLHFDQYREQLRARVTSEVVLDKLHGQVLS
                     QEQYERVLAENTRSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWEKGSKKG
                     LLPLSS"
BASE COUNT     1018 a   1202 c   1208 g    901 t
ORIGIN
```

Query Match   72.6%;   Score 3952;   DB 6;   Length 4329;
Best Local Similarity   94.9%;   Pred. No. 0;
Matches 4197;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 2;

```
Qy      1 ATGGCTGGGCGAGCCTGGGGCCGCGTTGGCCTTCTTTACTTGAGTTCTTCGTGAAGAAGAGGAGGAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGCTGGGCGAGCCTGGGGCCGCGTTGGCCTTCTTTACTTGAGTTCTTCGTGAAGAAGAGGAGGAG   60

Qy     61 CTGAAGGAGTTCCAGCTTCTTCTGCTCGGCCAATAAAAGCGCACTCCAGGAGCTCTTCGGGGTGAG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGAAGGAGTTCCAGCTTCTTCTGCTCGGCCAATAAAAGCGCACTCCAGGAGCTCTTCGGGGTGAG  120

Qy    121 ACACCCGCTCAGCCACAGAAGACGAGTGGCATGGGAGGTGGCCTGGCCCTCGTACCTGGTGGCTCAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACACCCGCTCAGCCACAGAAGACGAGTGGCATGGGAGGTGGCCTGGCCCTCGTACCTGGTGGCTCAG  180

Qy    181 TATGGGGAGCAGCGGGCGGGCCTGGGACCTAGCCCTTCCATCACCTGGGAGCAGAGATGGGGCTGAGG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATGGGGAGCAGCGGGCGGGCCTGGGACCTAGCCCTTCCATCACCTGGGAGCAGAGATGGGGCTGAGG  240

Qy    241 TCACTGTGGCGCCCAAGCCCAGGCAGGCCCACTCCATTCCCGTCATTCCCTACCTACAGCCA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACTGTGGCGCCCAAGCCCAGGCAGGCCCACTCCATTCCCGTCATTCCCTACCTACAGCCA  300

Qy    301 AGTGAACCCCACCTGGGGTCTCCCCAGCCAACCACCTCCACCCGCAGTGCTAATGCCCTGG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGTGAACCCCACCTGGGGTCTCCCCAGCCAACCACCTCCACCCGCAGTGCTAATGCCCTGG  360

Qy    361 ATCCATGAAATTGCCCGGCGGGGTGCACCCAGGGCTCAGGAGAAATCTGCCTCAGAGAGAAGGGGTTTTGAGACCAGCTG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCCATGAAATTGCCCGGCGGGGTGCACCCAGGGCTCAGGAGAAATCTGCCTCAGAGAGAAGGGGTTTTGAGACCAGCTG  420

Qy    421 CCTGACACATCTGGACGCCGCTGGGAGAGAAAATCTCTGCCTCACTCCCTACCCACAGCTCTT  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCTGACACATCTGGACGCCGCTGGGAGAGAAAATCTCTGCCTCACTCCCTACCCACAGCTCTT  480
```

(... alignment continues Qy/Db through to position 4332 ...)

```
Qy   4291 GAGAAGGGGCAGCAAAAAAGGGACTCTCTGCCACTCAGCAGCTGA  4332
          |||||||||||||||||||||||||||||||||||||||||||||
Db   4771 GAGAAGGGGCAGCAAAAAAGGGACTCTCTGCCACTCAGCAGCTGA  4812
```

Left column:

```
Db    481  CCAAGCTCCCAGACCATGAGTCTTCCAAGCCAGGAGTCACCCAAGCCCCCACATCCACA   540

Qy   1063  GCAGTGCTGGGGAGCTGGGGGATCCCCAGCCTCAGCCCAGCCTCAGAGAGCAGGAG    1122
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GCAGTGCTGGGGAGCTGGGGGATCCCCAGCCTCAGCCCAGCCTCAGAGAGCAGGAG     600

Qy   1123  GCTCCCTGGGACCCCAATGGCCTCTGGGATGAAACGTCAGGAATTCACACAGAAATCAGA  1182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GCTCCCTGGGACCCCAATGGCCTCTGGGATGAAACGTCAGGAATTCACACAGAAATCAGA  660

Qy   1183  GAAAGAGAGAGAGAGAAATCAGAGAAAGGCAGCAGCCCCCCCATGGGCAGCGGCGG     1242
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GAAAGAGAGAGAGAGAAATCAGAGAAAGGCAGCAGCCCCCCCATGGGCAGCGGCGG    720

Qy   1243  CCCCCACAGGGCGGCACCAGCCTACAGCCCCCACCACCTACCCCCATGGGAGCCTTCTGTGAGA  1302
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CCCCCACAGGGCGGCACCAGCCTACAGCCCCCACCACCTACCCCCATGGGAGCCTTCTGTGAGA  780

Qy   1303  GAGAGCCTCTGTTCCACATGGGCCTGGGAAAATTTTAACCAAAAAAATTCACACAG    1362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GAGAGCCTCTGTTCCACATGGGCCTGGGAAAATTTTAACCAAAAAAATTCACACAG   840

Qy   1363  CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGG  1422
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CTGCTACTTCTACAAAGACCTCACCCCAGAAGCCAAGATCCCCTGGTCAAGAGAAGCTGG  900

Qy   1423  CCTGATTATGTGGAAGTCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA   1482
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CCTGATTATGTGGAAGTCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA   960

Qy   1483  GGCCTGGATACCCAAGAACCTCGCATAGTCATATCTGCAGGGGCTGCTGGGGAAG  1542
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GGCCTGGATACCCAAGAACCTCGCATAGTCATATCTGCAGGGGCTGCTGGGGAAG  1020

Qy   1543  TCAAACACTGGGCCAGGCCAGGTGAAGGAAGCCTGGGGAGGAGGCCAAGAGGTCCCTCCAC  1602
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  TCAACACTGGGCCAGGCCAGGTGAAGGAAGCCTGGGGAGGAGGCCAAGAGGTCCCTCCAC  1080

Qy   1603  TTCCAGCAATGTCTTCTCATTCAGCTGCAGAGACTGGCCCAGTCCCAAGGTGGTGAGTCTC  1662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TTCCAGCAATGTCTTCTCATTCAGCTGCAGAGACTGGCCCAGTCCCAAGGTGGTGAGTCTC  1140

Qy   1663  GCTGAGCTCATCCGGAAAAAGATGGGACAGCCACCTCCGGGCTCCCAATTAGACAGATCCTGTCT  1722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GCTGAGCTCATCCGGAAAAAGATGGGACAGCCACCTCCGGGCTCCCAATTAGACAGATCCTGTCT  1200

Qy   1723  AGGCCAGAGGCGGCTGCTCTTCATCCTCCGATGGTGTAGATGAGTGAGGCCAGGAGGTCTTGCAG  1782
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  AGGCCAGAGGCGGCTGCTCTTCATCCTCGATGGTGTAGATGAGCCAGGAGGTCTTGCAG  1260

Qy   1783  GAGCCGAGTTCTGAGCTGTGTCTGGCACTGGAGCCGGCCGGAGGATGGGGTCTTGCTG  1842
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  GAGCCGAGTTCTGAGCTCTGTCTGGCACTGGAGCCGGCCGGAGGATGCACTGCTG  1320

Qy   1843  GGCAGTTTGCTGGGGGAAAACTATACTTCCCGAGGCATCCTTCCTGATCACGGCTCGGACC  1902
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  GGCAGTTTGCTGGGGGAAAACTATACTTCCCGAGGCATCCTTCCTGATCACGGCTCGGACC  1380

Qy   1903  ACAGCTCTGCAGAAACCTCATTCCTTCTTTGGACCAGCACAGTTGGGAGGAGGTCCCTGGG  1962
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  ACAGCTCTGCAGAAACCTCATTCCTTCTTTGGACGCAGCACAGTTGGGAGGAGGTCCCTGGG  1440

Qy   1963  TTCTCTGAGTCCAGCAGGAAGGAAATATTTCTACAGATATTTCACAGATGAAAAGGCAAGCA  2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  TTCTCTGAGTCCAGCAGGAAGGAAATATTTCTACAGATATTTCACAGATGAAAAGGCAAGCA  1500

Qy   2023  ATTAGAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGAGCTCTGCCCCTGTGTCTTGTGCCC  2082
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  ATTAGAGAGCCTTTAGGTTGGTCAAATCAAACAAAGAGAGCTCTGCCCCTGTGTCTTGTGCCC  1560

Qy   2083  TGGGTGTGTCCTGGCTGGCCTGCACTTGCCTGATGCAGCAGCAGCAGATGAAGCGAAGGAAAAACTC  2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  TGGGTGTGTCCTGGCTGGCCTGCACTTGCCTGATGCAGCAGCAGCAGATGAAGCGAAGGAAAAACTC  1620
```

Right column:

```
Qy   2143  ACACTGACTTCCAAGACCACCACCACAACCCTCTTGTCTACATTACCTTGCCCAGGCTCTCCAA  2202
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  ACACTGACTTCCAAGACCACCACCAACCCTCTTGTCTACATTACCTTGCCCAGGCTCTCCAA  1680

Qy   2203  GCTCAGCCATTGGGACCCCAGTCTCAGAGGACCCTCTGCTCTTCTGGCCTGTCTGAGCGGCATCTGG  2262
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  GCTCAGCCATTGGGACCCCAGTCTCAGAGGACCCTCTGCTCTTCTGGCCTGTCTGAGCGGCATCTGG  1740

Qy   2263  CAAAAAAAAGAGACCCCTTTTCAGTCAGATGAACCTCAGGAGGCAATGGGCCATC  2322
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1741  ATCTTCCACCTTCTTGAAGATGGGGTATTCTTCAAGAGCACCCATCCCCTCGACTACAGC  1800

Qy   2323  ATCTTCCACCTTCTTGAAGATGGGGTATTCTTCAAGAGCACCCATCCCCTCGACTACAGC  2382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  ATCTTCCACCTTCTTGAAGATGGGGTATTCTTCAAGAGCACCCATCCCCTCGACTACAGC  1860

Qy   2383  TTCATTCACCTCTGTTTCCCAAGAGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGCTCTG  2442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  TTCATTCACCTCTGTTTCCCAAGAGTTCTTTGCAGCAATGTCCTATGTCTTGGAGGCTGAG  1920

Qy   2443  AAGGGGAGGAGGGTAAAACATTCTAATTGCATCAACACACAGTTTCATCATAGATTTGGAAAAAGACCGGTTAGAGGCATAT  2502
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  AAGGGGAGGAGGGTAAAACATTCTAATTGCATCAACACACAGTTTCATCATAGATATTGGAAAAAGACCGGTTAGAGGCATAT  1980

Qy   2503  GGAATACAATGGGCCTGTTTGGGGGCATCAACCACACAGTTTCCTAAATTGGGCCTGTTAAGTGAT  2562
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  GGAATACAATGGGCCTGTTTGGGGGCATCAACCACACAGTTTCCTATTGGGCCTGTTAAGTGAT  2040

Qy   2563  GAGGGGGAGAGAGAGGATGGGAGAACATCTTTCACTGCCGGGCTGTCTCAGGGGAACCTG  2622
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  GAGGGGGAGGAGGAGGATGAGGAACATCTTTCACTGCCGGGCTGTCTCAGGGGAACCTG  2100

Qy   2743  ATGGGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGTGCACTTTCTGCATTAAAATTC  2802
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  TGCTTGTACGGAGACTCGGAACAAAAACGTTCCTGACACAAGTGATGGGCCCATTCGAAGA  2220

Qy   2683  TGCTTGTACGGAGACTCGGAACAAAAACGTTCCTGACACAAGTGATGGGCCCATTCGAAGA  2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  ATGGGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGTGCACTTTCTGCATTAAAATTC  2280

Qy   2803  AGCGCGCCAGTGAAGAAGCTTCAGCTGATTGAGGGGCAGGCAGGCGCACAGATCAACATGGAGC  2862
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  AGCGCGCCAGTGAAGAAGCTTCAGCTGATTGAGGGGCAGGCAGGCGCACAGATCAACATGGAGC  2340

Qy   2623  ATGCAGTGGGGTCCGTGCCTGCGCTGCAGCCACACTGTCTCCTCCAC  2682
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  ATGCAGTGGGGTCCGTGCCTGCGCTGCAGCCACACTGTCTCCTCCAC  2160

Qy   2923  TTCTCCGTCCTCAAAGGTCACCAGAGAAAACCTGAAGGAGCTTAAGTGGAAAACTCGCTG  2982
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  CCCACCATGGTAGTCCTGTTCAGGTGGCCCAGTCACAGATGCCTATTGGCAGATTCTC  2400

Qy   2983  AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGACGGCCCTGCCTGG  3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACGCCCTGAGACGGCCCTGCCTGG  2520

Qy   2461  ACCCTGCGGTTGCTGGCCTGTGTGCCTCACAGCTGAGGAACTGCAAGGACCTTGCCTTTGGG  2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  ACCCTGCGGTTGCTGGCCTGTGTGCCTCACAGCTGAGGAACTGCAAGGACCTTGCCTTTGGG  2580

Qy   3103  CTGAGGAGCCAACCAGAACCCTGACCGAGCGTGGACCTTCAATGTTGCTCACGGATGCT  3162
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  CTGAGGAGCCAACCAGAACCCTGACCGAGCGTGGACCTTCAATGTTGCTCACGGATGCT  2640

Qy   3163  GGAGCCAAAACACCTTTGCCAGAGAGACTGAGACAGCAGCCGCAAGCTACAGCGACTGCAG  3222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  GGAGCCAAAACACCTTTGCCAGAGAGACTGAGACAGCAGCCGCAAGCTACAGCGACTGCAG  2700
```

```
Qy    3223 CTGGTCAGCTGTGTGGCCTCACGTCGTCTGACTGGTGCCAGGAGGACCTGGCCTCTCGTGTGCTTAGTCGCC 3282
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CTGGTCAGCTGTGTGGCCTCACGTCGTCTGACTGGTGCCAGGAGGACCTGGCCTCTCGTGTGCTTAGTCGCC 2760

Qy    3283 AGCCCCAGCCTGAAGGAGCTAGAACTGCAGCAGCAACCTGGATGACGTTGGCGTGGCA 3342
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 AGCCCCAGCCTGAAGGAGCTAGAACTGCAGCAGCAACCTGGATGACGTTGGCGTGGA 2820

Qy    3343 CTGCTCTGTTGAGGGGCTCAGGCATCCTGCAAAACTCATACGGCCTGGGGTGGACCAG 3402
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 CTGCTCTTGTGAGGGGCTCAGGCATCCTGCAAAACTCATACGGCCTGGGGTGGACCAG 2867

Qy    3403 ACAACTCTGAGTGAATGAGGAGGCAGGAGAACTGAGGGGCCCTGAGCGCAGGAGAGAAACCTCAG 3462
Db    2868 ------------------------------------------------------------------ 2867

Qy    3463 CTGCTCATCTTCAGCAGACGGGAAAACCAAAGTGTGGATGACGGGCCTTACTGAGGGGCCTGGGGATACG 3522
           |||||||||||||||||||
Db    2868 --------GAAACCAAAGTGTGATGATGACCCCTTACTGAGGGGCCTGGGGATACG 2907

Qy    3523 GGAGAGGATGAGGTAAATAGCACACATCCTCACTCAACGGCACAGACTCGGACATCAGAGAAGGGGCG 3582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2908 GGAGAGGATGAGGTAAATAGCACACATCCTCACTCAACGGCACAGACTCGGACATCAGAGAAGGGGCG 2967

Qy    3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTCAAGGCACAGCAAGACTCGGATCAGAGGATCTT 3642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2968 GCTTCCCATGTTGCTCAGGCTAATCTCAAAACTCCTCAAGGCACAGCAAGACTCGGATCAGAGGATCTT 3027

Qy    3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGGTAGTACCCGTCAAACTCCTCTTGTGGGGCCT 3702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3028 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTGCAACTCCTCTTGTGGGGCCT 3087

Qy    3703 TCTCCCTGCCTCTCAAGGGGGACACTGCAGTACGGAAGCCTTTGGGGACTGACGAGTGACTTCTGG 3762
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3088 TCTCCCTGCCTCTCAAGGGGGACACTGCAGTACGGAAGCCTTTGGGGACTGACGAGTGACTTCTGG 3147

Qy    3763 GGCCCCACGGGGCCTACTGTGTCCTGGTAGTTGACAAAAGAAAAGAAACTTGTACCGGAGTT 3822
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3148 GGCCCCACGGGGCCTGTGTGGCTACTGAGGTAGTTGACAAAAGAAAAGAAACTTGTACCGGAGTT 3207

Qy    3823 CACTTCCCTGTAGCTGGCTCCTACCGGTGCTCTGCTTTGTTGTTGATATGAGA 3882
           |||||||||||||||||||||||||||||||||||
Db    3208 CACTTCCCTGTAGCTGGCTCCTACCGGTGCTCTGCTTTGTTGTTGATATGAGA 3267

Qy    3883 GAAGCGGTGACCGTTGAGATTGAAATTCTGTGGGACCCAGTTCCTGGGTGAGATCAAC 3942
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3268 GAAGCGGTGACCGTTGAGATTCAAATTCTGTGGGACCAGTTCCTGGGTGAGATCAAC 3327

Qy    3943 CCACAGCACAGCTGGGATGGTGGCAGGGGCCTCTCCTGCGACATCAAGGCTGAGCCTGGGACT 4002
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3328 CCACAGCACAGCTGGGATGGTGGCAGGGGCCTCTCCTGCGACATCAAGGCTGAGCCTGGGACT 3387

Qy    4003 GTTGGAAGCTGTGCACCTCCCTCACTTTGTTGGTCTTCAAGGAGGGGCCATGGTAGGACACATCC 4062
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3388 GTTGGAAGCTGTGCAACCTCCCCTCACTTTGTGGCTCTCCAAGGGGGCCATGGTGACACATCC 3447

Qy    4063 CTGTTTCAAAATGGCCCACTTTAAAGAGGAGGGGGAATGCTCCTGAGAACCCAGCCCAGGGTG 4122
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    3448 CTGTTTCCAAATGGCCCACTTTAAAGAGGAGGGGGATGCTCCTGAGAACCCAGCCCAGGGTG 3507

Qy    4123 GAGCTGCATCACATAGTTCTGGAAAAACCCCAGCTTCTCCCCCCCTTTGGGAGGTCCTCCTGAAA 4182
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3508 GAGCTGCATCACATAGTTCTGGAAAAACCCCAGCTTCTCCCCCTTTGGGAGGTCCTCCTGAAA 3567

Qy    4183 ATGATCCATAATGCCCTGCGCTTCCACCTTCCAGCCTGTCTGTGTGGTTGGTTGCTTTACCACCGC 4242
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3568 ATGATCCATAATGCCCTGCGCTTCATTCCCGTCACCTGTGTGGTTGGTTGCTTTACCACCGC 3627

Qy    4243 GTCCATCCTTGAGGAAGTCACCTTCCACCTTCCACTTCCACTGACTGACTGCTTCCATTGGG 4302
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    3628 GTCCATCCTTGAGGAAGTCACCTTCCACTTCCACTGACTGACTGCTCCCATTCGG 3687

Qy    4303 -------------------------------------------------------------- 4302
Db    3688 AAGGCCCATAGATGATCTAGAAATGAAAATTCCCAGTTTGTGGGAATCCACAAGCCACCCCCG 3747

Qy    4303 -------------------------------------------------------------- 4302
Db    3748 CTGACCCCACTTTATATGGGGCTGTGTCGTTACACTGTGTCTGGGTCTTGGGGTTCAGGGATGCTG 3807

Qy    4303 ------AAGGACTGGAGCTCTGCTGTCTATCCAAGCCCCTGGAGAAGGACCAGCTGTTC 4350
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3808 GAAATACTCCCCAAGGAACTGGAGCTCTGCTATCCGAAGCCCTGGAGAAGGACCAGCTGTTC 3867

Qy    4351 TCGGAGTTCTTACGTTGGGCCACTTGGGATCAGGGGATCAGGGTGCAAGTGAAGATGAGAAA 4410
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3868 TCGGAGTTCTTACGTTGGGCCACTTGGGATCAGGGGATCAGGGTGCAAGTGAAGACAAGAAA 3927

Qy    4411 GATGAGACTCTGGTTGTGTGGGAGGGGCCTTGGTGAAACCAGGAGGATCTCATTGCCTCTT 4470
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3928 GATGAGACTCTGGTTGTGTGGGAGGGGCCTTGGTGAAACCAGGAGAGATCTCATTGCCTCTT 3987

Qy    4471 CTGATCCTTCCAGCCCGCCATAGCCGTACCTTCACCCTCTGGAATGCCCCGCAGTTGCTGCAC 4530
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3988 CTGATCCTTCCAGCCCGCCATAGCCGTACCTTCACCTCTGGAATGCCCCGCAGTTGCTGCAC 4047

Qy    4531 TTTGTTGGACCAGTATCTGAGAGCAGCTACTTGATGAGCCCAGTGACATCGGTGAGGGTCTTG 4590
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4048 TTTGTTGGACCAGTATCTGAGAGCAGCTACTTGATGAGCCCAGTGACATCGGTGAGGGTCTTG 4107

Qy    4591 GACAAAACTGCATGGAGACAGTGCTGAGCCAGTAGGAGGGTGCTGGCTGAGAAC 4650
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4108 GACAAAACTGCATGACACAGTGCTGAGCCAGTAGGAGAGGGTGCTGGCTGAGAAC 4167

Qy    4651 ACGAGGCCCAGCAGATGCCGAAGCTGTTCAGCTTGAGCTGCAGTCCTGGGGACCGGAAGTGC 4710
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4168 ACGAGGCCCAGCAGATGCCGAAGCTGTTCAGCTTGAGCTGCAGTCCTGGGGACCGGAAGTGC 4227

Qy    4711 AAAGATGGACTCTACCAACCCCTGAAAGAGACCCCATCCCACCTCATTATTGGAACTCTGG 4770
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4228 AAAGATGGACTCTACCAAGCCCTGAAAGGAGACCCCATCCTCACCTTATTGGAACTCTGG 4287

Qy    4771 GAGAAGGGGCAGCAGCAAAAAAGGGACTCCTGCCACTCAGCAGCTGA 4812
           |||||||||||||||||||||||||||||||||||||||||
Db    4288 GAGAAGGGGCAGCAGCAAAAAGGGACTCCTGCCACTCAGCAGCTGA 4329
```

RESULT 14
HSM800983

```
LOCUS       HSM800983               4123 bp    mRNA            linear   PRI 18-FEB-2000
DEFINITION  Homo sapiens mRNA; cDNA DKFZp586O1822 (from clone DKFZp586O1822);
            partial cds.
ACCESSION   AL117470
VERSION     AL117470.1  GI:5911938
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4123)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMPZ [Biomedical Research Center at the Charite,
            Berlin/Germany] within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp586O1822) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..4123
```

```
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp586O1822"
                /tissue_type="uterus"
                /clone_lib="586 (synonym: hutel). Vector pSport1; host
                DH10B; sites NocI + SalI/MluI"
                /dev_stage="adult"
     gene        1..4123
                /gene="DKFZp586O1822"
     CDS         <1..3580
                /gene="DKFZp586O1822"
                /note="KIAA0926, partly, alternative spliced"
                /codon_start=2
                /product="hypothetical protein"
                /protein_id="CAB55945.1"
                /db_xref="GI:5911939"
                /translation="LLLQRPHPRSQDPLVKRSWPDVVEENRGHLIEIRDLFGPGLDTQ
                EPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAEL
                IGKDGTATPAPIROILSRPERLLFILDGVDEPGWVLQEPSSELCLHWSQPQPADALLG
                SLIGKTILPEASFLITARTTALQNIIPSLEQARWVEVLGFSESSRKEYFRYFDERG
                AIRARRLVKSNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTTLCLYKLAQ
                ALQAQPLPQLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPI
                PLSYSFIHLCPEFFAAMSTVLEDKRGRGKHSNCIIDLEKTLEAVGIHGLFGASTTRF
                LLGLLSDEGGREMENTFHCRLSQGRNLMQWVPSLLQLILQPHSLESLHCLYETRNKTFL
                TQVMAHFEEMGMCVETDMELLVCTFCIKFSRHVKKLQLIEGRQHRSTWSPTMVVLFRW
                VPVDAYWQILPSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRPRCLLSTLRLAGC
                GLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLCORLROPSCKLQRLQLVSCG
                LTSDCCQDLASVLSASPSLKELDLQQNNLDDVGVRLLCEGLRHPACKLIRLGLDQTTL
                SDEMRQELRALEQEPQLLIFSRRKRPSVMTPTEGLDTGEMSNSTSSLKRQRLGSERAA
                SHVAQANLKLLDVSKLFPIAEIAESSPEVVPELLCVPSPASQGDLHTKPLGTDDDF
                WGPTGPVATEVVDKEKNLRYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLG
                EINPQHSWMWAGPLLDIKAPEGAVEAVHLPHPVALQGGHVDTSLFQVAHFKEEGMLLE
                KPARVELHHIVLENPSFSPLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFHLYLI
                PSDCSIRKAIDDLEMKFQPVRIHKPPPLTPLYMGCRYTVSGSSGSGMLEILPKELELCY
                RSPGEDQLFSEFYVGHLGSGYHLQVDKDQETLVWEALVRVGDLMPATTLIPPARIAV
                PSPLDAPQLLHFVDQYREQLIARVISVEVVLDKLHGQVLSQEQYERVLAENTRPSQMR
                KLFSLSQSWDRKCDGLYQALKETHPHLIMELWEKGSKXGLLPLSS"
     polyA_signal  4057..4062
                /gene="DKFZp586O1822"
     polyA_site    4081
                /gene="DKFZp586O1822"
     BASE COUNT  1002 a  1086 c  1136 g  899 t
     ORIGIN
```

Query Match   69.9%;   Score 3807.8;   DB 9;   Length 4123;
Best Local Similarity   96.7%;   Pred. No. 0;
Matches 3951;   Conservative   0;   Mismatches   2;   Indels 132;   Gaps   1;

```
Qy     1365 GCTACTTCTACAAAGACCTCACCCCAGAGAGCCAAGATCCCTGGTCAAGAGAAGCTGGCC 1424
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GCTACTTCTACAAAGACCTCACCCCAGAGAGCCAAGATCCCTGGTCAAGAGAAGCTGGCC   60

Qy     1425 TGATTATGTGGAGGAGAATCGAGGACAATTTAATTGAGATCAGAGACTTATTTGGGCCAG 1484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 TGATTATGTGGAGGAGAATCGAGGACAATTTAATTGAGATCAGAGACTTATTTGGGCCAG  120

Qy     1485 CCTGGGATACCCAAGAAACCTCGCATAGTCATACTCTGCAGGGGGCTTGCTGGGAAGTC 1544
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CCTGGGATACCCAAGAAACCTCGCATAGTCATACTCTGCAGGGGGCTTGCTGGGAAGTC  180

Qy     1545 AACACTGGCCAGGCCAGGTGAAGGAAGCCTGGGGGGAGGAGGCCAGCTGTAGTGGGGAGGGACCGGCTT 1604
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AACACTGGCCAGGCCAGGTGAAGGAAGCCTGGGGGGAGGAGGCCAGCTGTAGTGGGGAGGGACCGGCTT  240

Qy     1605 CCAGCAATGTGTTCTTCACTTCAGCTGTGCAGAGGAGCTGGCCCAGTCCAAAGTTGGTGAGTGAGTCTCGCC 1664
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CCAGCATGTGTTCTTCACTTCAGCTGTGCAGAGAGCTGGCCCAGTCCAAAGTGGTTGAGTGAGTCTCGCC  300

Qy     1665 TGAGCTCATCGGAAAAGATGGGAGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCTCTAG 1724
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TGAGCTCATCGGAAAAGATGGGAGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCTCTAG  360

Qy     1725 GCCAGAGAGCGGGCTGCTCTTCATCCTTCGGATGGGTTGTAGATGAGCCAGGATGGGTTCTTTGCAGGA 1784
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCAGAGAGCGGGCTGCTCTTCATCCTTCGGATGGGTTGTAGATGAGCCAGGATGGGTTCTTTGCAGGA  420

Qy     1785 GCCGGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACCAGCGGACCGGCGGATGGGTTCTTTTGGGGTT 1844
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GCCGGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACCAGCGGACCGGCGGATGGGTTCTTTTGGGGTT  480

Qy     1845 CAGTTTGCTGGGGGAAAAACTATTACTTTCCCGGAGGGCATCCTTCCTTGATCACGGGCTGGGACCAC 1904
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CAGTTTGCTGGGGGAAAAACTATTACTTTCCCGGAGGGCATCCTTCCTTGATCACGGGCTGGGACCAC  540

Qy     1905 AGCTCTGCAGAACCTCATTCCTTCTTGTGGGGGAAAAACTATACTTTCACAGATACATGTGGGGTT 1964
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 AGCTCTGCAGAACCTCATTCCTTCTTTTGGGGGAAAAACTATAACTTTCACAGATACATGTGGGGTT  600

Qy     1965 CTCTGAGTCCAGCAGGAGGAAGAATATTTCTACAGATATTTCACAGATATATTCACAGCAAT 2024
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTCTGAGTCCAGCAGGAGGAAGAATATTTCTACAGATATTTCACAGATATATTCACAGCAAT  660

Qy     2025 TACAGACCTTAGTTTGGTCAAATCAAACAAAAGAGAGCTCTGGGGCCCCTGTTCTCTGTTGTGCCCTG 2084
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TACAGACCTTAGTTTGGTCAAATCAAACAAAAGAGAGCTCTGGGGCCCCTGTTCTCTGTTGTGCCCTG  720

Qy     2085 GGTGTCCTGGCCTGGCCTGCACTTGCCTGAATGCAGCAGTGAAGCGGGAAGGAAGAAAACTAC 2144
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GGTGTCCTGGCCTGGCCTGCACTTGCCTGATGCAGCAGTGAAGCGGGAAGGAAGAAAAACTAC  780

Qy     2145 ACTGACTTCCAAGACCACCACAACCCTCTGTCTCTGTACATTACCTTGGCCCAGGCCTCCCAAGC 2204
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 ACTGACTTCCAAGACCACCACAACCCTCTGTCTCTGTACATTACCTTGGCCCAGGCCTCCCAAGC  840

Qy     2205 TCAGCCATTGGGACCCCAGCTCAGAGACCTCAGGAAGCATGGGTTAGATGGGGCCCATCAT 2264
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TCAGCCATTGGGACCCCAGCTCAGAGACCTCAGGAAGCATGGGTTAGATGGGGCCCATCAT  900

Qy     2265 AAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCAT 2324
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 AAAAAAGACCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCAT  960

Qy     2325 CTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGACCACCCATCCCTCTGAGCTACAGCTT 2384
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGACCACCCATCCCTCTGAGCTACAGCTT 1020

Qy     2385 CATTCACCTCTGTTTCCAAGAGAGTTCTTTGCAGCAAATGTCCTATGTCTGGGAGGATGAGAA 2444
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CATTCACCTCTGTTTCCAAGAGAGTTCTTTGCAGCAAATGTCCTATGTCTGGGAGGATGAGAA 1080

Qy     2445 GGGGAGGAGGTAAACATTCTAATTGCATCATAGATTTGAAAAAGACGCTAGAAGCATATGG 2504
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GGGGAGGAGGTAAACATTCTAATTGCATCATAGATTTGAAAAAGACGCTAGAAGCATATGG 1140

Qy     2505 AATACATGGCCTGTTTGGGGCATCAACCACACGTTTCCTATTGGGGCCCTGTTAAGTGATGA 2564
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AATACATGGCCTGTTTGGGGCATCAACCACACGTTTCCTATTGGGGCCCTGTTAAGTGATGA 1200

Qy     2565 GGGGGAGGAGGACGATGGAGAGAACATCTTTCACTCCCGGCGTCTCAGGGAGGGGAACCTGAT 2624
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 GGGGGAGGAGGATGGAGAGAACATCTTTCACTCCCGGCGTCTCAGGGAGGGGAACCTGAT 1260

Qy     2625 GCAGTGGGGTGTCCCCGTCCCTGCAGTCGTCTGCGAGCCCACACTCTCTCGAGTCCCTCCACTG 2684
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GCAGTGGGGTCCCCGTCCCTGCAGTCGTCTGCGAGCCCACACTCTCTCGAGTCCCTCCACTG 1320

Qy     2685 CTTGTACGAGACTCGGACAACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAAAT 2744
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CTTGTACGAGACTCGGACAACAAAACGTTCCTGACACAAGTGATGGCCCATTTCGAAGAAAT 1380

Qy     2745 GGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGTACTTTCTGCATTAAAATTCAG 2804
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 GGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGTACTTTCTGCATTAAAATTCAG 1440

Qy     2805 CCCGCCACGTGAAGAGAAGCTTCAGCTGATTGAGTGAGGGCAGCCAGCACAGATCAACATGGAGCCC 2864
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 GGGCATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGTACTTTCTGCATTAAAATTCAG 1440
```

RESULT 15

LOCUS   AX089769   4556 bp   DNA   linear   PAT 21-MAR-2001
DEFINITION   Sequence 9 from Patent WO0116170.
ACCESSION   AX089769
VERSION   AX089769.1 GI:13443941
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Reed,J.C.
  TITLE   Card proteins involved in cell death regulation
  JOURNAL   Patent: WO 0116170-A 9 08-MAR-2001;
             The Burnham Institute (US)
FEATURES   Location/Qualifiers
     source   1..4556
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS   1..4365
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC34903.1"
                     /db_xref="GI:13443942"
                     /note="unnamed protein product"
                     /translation="MAGGAWRLACYLEFLKKEELKEFQLLLANKAHSRSSSGETPAQ
                     PEKTSGMEVASYLVAQYEGRQRAWDLALHTWEQMGLRSLCAQAQEGAGHSFSPYSPSE
                     PHLGPSPQFTSTAVLMPWIHELPAGCTQGSERRVLRQLPDTSGRWREISASLLYQAL
                     PSSPDHESPSQESPNAPTSTAVLGSWSGPPOPSLAPREQEAPGTQWPLDETSGIYYTE
                     IREREREKSEKGRPWAAVVGTPPQAHTSLQPHHHPWEPSVRESLCSTWPWKNEDFNQ
                     KFTQLLLLQRPHPRSQDPLVRRSWPDYVEENRGHLIEIRDLFPGPGLDTQEPRIVILQG
                     AAGIGKSTLARQVKEAWGRGQLYGDRFQHVFYFSCRELAQSKVVSLAELIGKDGTATP
                     APIRQILSRPERLLFIIDGVDEPGWVLQEPSSELCLHWSQPQPADALLGSLLGKTILP
                     EASFLITARTALQNLIPSLEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRLVK
                     SNKELWALCLVPWVSWLACTCLMQQMRKEKLTLTSKTTTLCLHYLAQALQAQPLGP
                     QLRDLCSLAAEGIWQKKTLFSPDDLRKHGLDGAIISTFLKMGILQEHPIPLSYSFIHL
                     CFQEFPAAMSVYLEDEXGRGKHSNCIIDLEKTLEAYGIHGLFGASTTRFLLGLLSDBG
                     EREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESLHCLYETRNKTFLTQVMAHFEE

BASE COUNT   1100 a   1223 c   1256 g   977 t
ORIGIN

Query Match   62.5%; Score 3400.4; DB 6; Length 4556;
Best Local Similarity 93.7%;
Matches 3555; Conservative   0; Mismatches 236; Indels   3; Gaps   1;

Qy   523 ATGGCTGGGCCGGAGGCCTGGGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 582
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCTGGGCCGGAGGCCTGGGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAAGGAGGAG 60

Qy   583 CTGAAAGGAGTTCAGCTTCTGCTTGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   61 CTGAAAGGAGTTCAGCTTCTGCTTGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120

Qy   643 ACACCCGCTCAGCCCAGAGAAGACAGAGTGGCATGGAAGGTGGCCTCGGTACCTGGTCAG 702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACACCCGCTCAGCCCAGAGAAGACAGAGTGGCATGGAAGGTGGCCTCGGTACCTGGTCAG 180

Qy   703 TATGGGGGACCAGCGGGCGGGGGGGGGGGTGGGGACCTAGCCCTTGGGGACTGGGCAGAGG 762
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   181 TATGGGGAGCAGCGGGCGGGGGGGGTGGGGACCTAGCCCTTGGGGACTGGGCAGAGG 240

Qy   763 TCACTGTGCGGCCCAAGCCCAAGGAAGGGGCGCCAGCCCACTCTCCCTGGGGACCGCCA 822
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCACTGTGCGGCCCAAGCCCAAGGAAGGGGCGCCAGCCCACTCTCCCTGGGGACCGCCA 300

Qy   823 AGTGAAACCCACCTGGGGTCTCCCAGCCAAACCACCTCCACCCGCGCAGTGCTAATGCCTGG 882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGTGAAACCCACCTGGGGTCTCCCAGCCAAACCACCTCCACCCGCGCAGTGCTAATGCCTGG 360

Qy   883 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGCGCCCACCCACATCCACA 942
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATCCATGAATTGCCGGCGGTGCACCCAGGGGTCACCCAGAGAGAAGGGTTTTGAGACAGCTG 420

Qy   943 CCTGACACATCTCGACGCCGCTGGAGGAGAAAATCTCTGCCTCACTCCCTCCACCAAGCTCTT 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCTGACACATCTGGACGCCGCTGGAGGAGAAAATCTCTGCCTCACTCCCTCCACCAAGCTCTT 480

Qy   1003 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGCGCCCCACCAATCCACA 1062
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCAAGCTCCCCAGACCATGAGTCTCCAAGCCAGGAGTCACCCAAGCGCCCCACCCACATCCACA 540

Qy   1063 GCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCAGCCTAGCACCCCAGAGAGCAGGAG 1122
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCAGTGCTGGGGAGCTGGGGATCCCCACCTCAGCCCAGCCTAGCACCCCAGAGAGCAGGAG 600

Qy   1123 GCTCCCTGGGACCCCAAATGCCCTCTTGGATGAGAACGTCAGGAATTTACTACACAGAAATCAGA 1182
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCCTGGGACCCCAAATGCCCTCTTGGATGAGAACGTCAGGAATTTACTACACAGAAATCAGA 660

Qy   1183 GAAAGAGAGAGAGAAATCAGAGAAAGGCAGGCGCCCCCATGGGCGCAGCGGCAGGGCTAGGGAACG 1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GAAAGAGAGAGAGAAATCAGAGAAAGGCAGGCGCCCCCATGGGCGCAGCGGTGGTAGGGACG 720

Qy   1243 CCCCCACAGGGCGCACACCAGCCTACAGCCCCCACCACCAGCCATGGGAGCCTTCGTGAGA 1302
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCCCACAGGGCGCACACCAGCCTACAGCCCCCACCACCAGCCATGGGAGCCTTCGTGAGA 780

Qy   1303 GAGAGCCCTCGTGTTCCACATGGCCCTGGAAAAAATGAGGATTTTAACCAAAAATTCACACAG 1362

═══════════════════════════════════════

MGMCVETDMELLVCTFCIKFSRHVKKLLQLIEGRQHRSTWSSPTMVVLFRWVPVTDAYWQ
ILFSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETFLRLAGCGLTAEDCKD
LAFGLRANQTITELDLSFNVLTDAGAHLCQRLRQPSCKLQLVSCGLTSDCCCDL
ASVLSASPSLKELDLDQNNLDDVGVRLLCEGLRHPACKLRLGLDQTTLSDEMRQELR
ALEQEKFQLLIFSRRKPSVMTPTEGLLDTGEMSNSTSSLKRQRLGSSERAASHVAQANLX
LLDVSKIFPIAEIAEBSSSPEVVPVELLCVPSPASQGDLHTKFLGTDDFLGPEGNVDV
ELIDKSTNRYSVWPPFTRAGWYLWSATGLGFLVRDEVTVTIAFGSSWSQHLALDLQHHEQW
LVGGPLFDVTAEPEEAVAEIHLPHFISLQGEVDVSWFLVAHFKNEGWVLEHPARVEPF
AIDDEEDRFHGVRLLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHFS
KFYAGQMKEPIQLEITEKRHGTLVWDTEVRPVDLQLVAASAPPFSGGAAFVKENHRQL
QARMGDLKGVLDDLQDNEVLTENEKELVEQEKTRQSKNEALLSMVEKKGDLALDVLFR
SISERDDYLVSYLRQQNL"

Db   781 GAGAGCCTCTGTTCCACATGGCCCTGGGAAAAATGAGGATTTAAACAAAAAATTCACACAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1363 CTGCTACTTCTACAAAAGAGACTCACCCAGAAGCCAAGATCCCCTGGGTCAAGAGAAGCTGG 1422
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGCTACTTCTACAAAAGAGACTCACCCAGAAGCCAAGATCCCCTGGGTCAAGAGAAGCTGG 900

Qy 1423 CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGGCCCA 1482
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CCTGATTATGTGGAGGAGAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGGCCCA 960

[... sequence alignment continues, Db/Qy paired rows ...]

```
Qy  3523 GGAGAGATGAGTAAATAGGACATCCTCACTCACTCAAGGGCAGAGAGACTCGGATCAGAGAGGGCG 3582
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 GGAGAGATGAGTAAATAGCACATCCTCACTCACGGGCAGAGAGACTCGGATCAGAGAGGGCG 3060

Qy  3583 GCTTCCCAATGTTGCTCAGGCTAATCTCAAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3642
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 GCTTCCCAATGTTGCTCAGGCTAATCTCAAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3120

Qy  3643 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGAAACTCTTGTCGGTGCCT 3702
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 GCTGAGATTGCAGAGGAAAGCTCCCCAGAGGTAGTACCGGTGAAACTCTTGTCGGTGCCT 3180

Qy  3703 TCTTCCTGCCTCTCAAGGGGGACCTGCATACGAAGGCCTTTGGGGACTGACGGATGACTTCTGG 3762
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 TCTTCCTGCCTCTCAAGGGGGACCTGCATACGAAGGCCTTTGGGGACTGACGGATGACTTTCTG 3240

Qy  3763 GGGCCCACGGGGGCCTGTGGCCTACTGGAGGTAGTTGACAAAGAAACTTGTACCGAGTT 3822
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GGGCCTGAAAGGAAAATGTGGATTGGAGTTGATTGATAAGAGCACAAAACAGAGATACAGCGTT 3300

Qy  3823 CACTTCCCTGTAGCTGGTCTCCTACCGGCTGGGCCCAACACGGGTGTCTCTGCTTTGATGATGAGA 3882
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 TGGTTCCCCACTGCTGGGCTGGGTTGTAGTCAGCCACAGGCCCTCGGGGTTCCTGGTAAGG 3360

Qy  3883 GAAGCCGTGACCGTTGAGATTGAAATTCTGTGTGTGGGGACCAGTTCCTGG---GTGAGATC 3939
         ||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db  3361 GATGAGGTCACAGTGACGATTGCGTTTGGTTCCTGGAGTCAGCGACCTGGCCCTGGGACCTG 3420

Qy  3940 AACCCACAGCACAGCTGGATGGTGGGCAGGGCCTTCTGCTGGACATCAAGGCTGAGCCTGGA 3999
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 CAGCACCAATGAACAGTGGCTGGTGGGGGGCCCCTTGTTTGANGTCACTGCAGAGCCAAGT 3480

Qy  4000 GCTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCCAAGGGGGGCCATGTGGACACA 4059
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 GAGGCTGGTCGGCCGGAAAATCCACCTTCCCCCACTTCATCCTCCTGGAGGTTGGGACGTC 3540

Qy  4060 TCCCTGTTCCAAATGGCCCACTTTAAAGGAGGAGGGGAGCTCCTGGAGAAGCCAGCCAGG 4119
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 TCCTGGTTTCTCGTTGCCCATTTTAAGGAATGAAGGATGGTCCTGGACCATCCAGCCCGG 3600

Qy  4120 GTGGGAGCTGCATCACATAGTTCTGGAAAAACCCCAGCGTTCTCCCCCTTGGGGAGTCCTCCTG 4179
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 GTGGGAGCCTTTCTATGCTGTCCTGGGAAAGCCCCAGCTTCTCTCTGATAGGGCATCCTGCTG 3660

Qy  4180 AAAAATGATCCATAAATGCCCTGCGCCTTCATTCCGTCACCTCTGTGGTGTTGCTTACCAC 4239
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 CGGATCGGCCAGTGGGATCGGCCCTTCTCCATCCCCATCACTTCAAACACATTGATCTATTAT 3720

Qy  4240 CGGCGTCCATCCTGAGGAAGTCACCTTCCCACCTTCCCAAGTGACTGCTCCATT 4299
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 CACCCCCCACCCCGAAGATAATTAAGTTCCACTTGTATACCTTGTCCCCAGCGACGGCCTTGCTA 3780

Qy  4300 CGGAAAGGAAACTGGA 4313
         ||||
Db  3781 ACAAAGGCGGATAGA 3794
```